THIS NUMBER COMPLETES VOLUME 12

Vol. 12 DECEMBER, 1941 No. 4

JAN 3:42 W

CHILD DEVELOPMENT

YEA

Editorial Board

CARROLL E. PALMER, *Chairman*
E. V. McCOLLUM
FREDERICK F. TISDALL
RICHARD E. SCAMMON
MANDEL SHERMAN
HAROLD E. JONES
F. STUART CHAPIN

PUBLISHED QUARTERLY BY THE SOCIETY FOR RESEARCH IN CHILD DEVELOPMENT
NATIONAL RESEARCH COUNCIL
2101 CONSTITUTION AVENUE
WASHINGTON, D. C.
Made in United States of America

CHILD DEVELOPMENT

Vol. 12

DECEMBER, 1941

No. 4

Editorial Board

E. V. McCOLLUM
Professor of Biochemistry, School of Hygiene and Public Health, The Johns Hopkins University.

RICHARD E. SCAMMON
Distinguished Service Professor Graduate School University of Minnesota

HAROLD E. JONES
Director, Institute of Child Welfare, and Professor of Psychology, University of California.

FREDERICK F. TISDALL
Director, Research Laboratories, Department of Pediatrics, University of Toronto, and Attending Physician, Hospital for Sick Children, Toronto.

MANDEL SHERMAN
The University of Chicago, Chicago, Illinois.

F. STUART CHAPIN
Professor and Chairman of Sociology and Director Graduate Social Work, University of Minnesota.

CARROLL E. PALMER, *Chairman*
Past Assistant Surgeon, U. S. Public Health Service.

PUBLISHED QUARTERLY BY THE SOCIETY FOR RESEARCH IN CHILD DEVELOPMENT
NATIONAL RESEARCH COUNCIL
2101 CONSTITUTION AVENUE
WASHINGTON, D. C.

Made in United States of America

POPULARITY AMONG PRESCHOOL CHILDREN[1]

ROSEMARY LIPPITT

INTRODUCTION

It is of interest to determine the concomitants of popularity among preschool children, for much of a child's adjustment in a group seems to depend upon the degree to which others find him acceptable. The good fellow, the leader, the isolate are partly defined by their relations to a particular group structure; at the same time there are some individuals who seem to make friends easily in any new group. Why is it that some children seem to be unpopular and unwanted members of a play group? Certainly preschool children experience acceptance and rejection by their associates, but some children continually appear to meet with rejection and others with acceptance. The following five minute record was taken from one of the groups of children used in this study and illustrates how a popular and unpopular child may be received by their associates. Child A was the most popular child in the group and Child X was the least popular. Children C, D, E, F and G were of the middle popular group.

> Children C, D and E are building a train of large box-sized blocks while carrying on a constant stream of conversation. Child X (unpopular) approaches within six feet of the "to be train". D stops work and with block in hand says to X, "You can't come over here. We don't like you." Then turning to C and E, "We don't want him, do we?" C and E hesitate a moment and respond, "No! We don't want him." D chimes in, "We don't like him." E yells out, "He is a poo-poo, we think he is a bad boy." All yell together, "He is a a poo-poo; he is a poo-poo." They laugh and C comes from behind the structure with a long block in his hand and flourishes it at X saying, "You get away from here you - you bad so-en-so. I'll smack you." X quiet until now, draws near the building and kicks at it. E rushes out to hit at X screaming, "Get away you bad - bad boy. We don't want you." C and D charge on X with blocks and drive him away, X leaves and approaches another group. C, D and E continue building the train and discussing how they dislike X. X approaches F and G and asks, "What you doing?" F responds, "She and me are playing house and we don't need a daddy." X kicks the table and the dishes that G had just placed there for a play meal fall to the floor. F rather angry hits at X and calls to the teacher, "Teacher, come quick! X is breaking into our house." X rushes into another room in a fit of temper. At this moment child A (popular) arrives upon the scene, and stands in the room a moment to get his bearings, much as X had done a few minutes earlier. A's eyes fall upon the train and at that moment E says, "A, you may play with us, can't he?" C to A, "See our fine train, A, it's almost ready to start. It's a stream-liner." The four children continue building.

[1]This is a condensation of a doctoral thesis carried on at the Iowa Child Welfare Research Station under the direction of Doctor Beth Wellman.

Of course, one cannot make all children very popular, but might not the scape-goat be abolished from our preschools by a better understanding of what seems to be characteristic of the children who are chosen popular and unpopular by their associates?

The present study has tried to determine the concomitants of popularity in preschool children with emphasis on the following items: 1) intensity of social participation; 2) directing, compliant, resistant, communicative, friendly, and hostile behavior; 3) attractiveness of personality; 4) physical attractiveness; 5) cooperation in the routines of the preschool; 6) mental age; 7) intelligence quotient; 8) chronological age; 9) the number of days of attendance at the preschool during the semester the data were gathered; 10) the number of observed minutes in which oral nervous habits were manifested; 11) gross and fine motor ability; 12) the reaction of the child when faced with failure; and 13) the constructive-destructiveness of the child as measured by Ackerman's test.

DESCRIPTION OF SUBJECTS

The subjects were 45 children attending the Iowa Child Welfare Research Station preschools during 1939 and 1940. Group A contained 15 subjects, 11 boys and 4 girls, who were members of a group of 22.[2] The mean intelligence quotient for this group was 133.3 with a range from 113 to 176,[3] and the chronologocal age ranged from 41 to 56 months with a mean age of 48.7 months.[4] Group B contained 9 subjects, 7 boys and 2 girls, who were members of a group of 13 children. The mean IQ of this group was 128.3 and ranged from 108 to 150. The age range was from 55 to 67 months with a mean age of 60.1 Group C contained 21 subjects, 11 boys and 10 girls, and was taken from a group of 22 children. The range in IQ for Group C was from 102 to 190 with a mean of 130.1, and the range in C.A. was from 38 to 49 months with a mean of 43.5 months. Ten of the Group A children remained in attendance at these preschools the following year and were available 9 months later for retesting. These 10 children comprised Group D; they were members of a preschool group of 26 which was divided into two subgroups of 13 in the school routine. There were 5 members of Group D in each subgroup. The mean IQ of Group D was 142.5 with a range from 111 to 176. The chronological age of the D-children ranged from 49 to 61 months with a mean of 56 months. The occupational level of the parents of these children was professional and semi-professional. A description of the preschool which these children attended, in terms of teacher's philosophy, curriculum, and teaching techniques, is given by Updegraff et al. (15).

GENERAL PROCEDURE

All children had been in attendance at the preschool at least six

[2] *Due to prolonged absences some of the children of each preschool group had to be eliminated from the study.*

[3] *Intelligence test scores were taken from tests administered within three months of the time of collection of data. (Revised Stanford-Binet Form L.)*

[4] *C.A. was computed from the mid-point of the time of collection of the data.*

weeks before the data of this investigation were gathered. Observations in the normal preschool situation were made first, and the testing program was begun after the children were familiar with the investigator.[5] The children were not taken in any particular order for testing, but were approached when the time seemed opportune, i.e. a child was approached when changing activity or standing idle. Results of a particular test were obtained on a group in as short a space of time as possible. In case of absences the measure was obtained soon after the child returned to preschool. This procedure was followed for all the measures except the intelligence and failure tests. The former was given on the day of the month, or thereabouts, on which the birthday fell.[6] The failure situation was given after the test of intelligence. No child was given more than two tests on any one day.

The tests were presented in the following order with few exceptions: 1) the Mosaic[7] test, which provided good introductory material, as the children were attracted by the brightly colored pieces; 2) the test of fine motor co-ordination; 3) the constructive-destructive situation; 4) popularity pairings; and 5) the gross motor items. Observations of social behavior and nervous habits were taken prior to the testing and continued to the end of the testing program. Groups A and B were tested in the spring of 1939 during the months of April and May. Data were gathered on groups C and D in the winter of 1939-40, from the last of November to the end of January.

The teacher estimates of child popularity and Merrill-Palmer ratings of compliance with routine, attractiveness of personality, and physical attractiveness were obtained at the midpoint of the study of each group.

TECHNIQUES USED

Popularity Preferences

The paired comparison method was employed in obtaining a popularity score for each member of a group from his immediate associates. This procedure necessitated each child giving his preference to one of the names in each pair with which he was presented. The list of pairs was continued until every name had been paired with every other name. This necessitated each child being presented with $\frac{n-1\ (n-2)}{2}$ pairings. The subject was not called upon to state preferences in pairs containing his own name.

Thurstone (18) justifies the use of the paired comparison method for describing objectively the attitudes of a group toward nationalities and races. In his article treating the application of the paired comparison method to social attitudes he states, "The same method can undoubtedly be used also for describing the attitudes of a group toward other social values which are saturated with prejudice and bias...and is free from

[5] *The present experimenter gave all the tests except the failure and intelligence.*

[6] *The tests of intelligence and failure were administered by trained psychometrists of the Iowa Child Welfare Research Station.*

[7] *The results of the Mosaic test have not yet been analysed. For description of this test see the Am. J. Psychiat., 9, 1939, 232-236.*

the effects of the personal bias of the investigator." (p. 423)

Koch (9) utilized the paired comparison method in obtaining popularity scores among preschool children. She presented the pairings so that each pair was proffered twice, the order of presentation of the names in a pair and the order of the pairs being reversed. By these means she hoped to prevent the serious marring of her results by a tendency on the part of the child to repeat the last names of a series. Thus each child was given n-1 (n-2) pairings, the reversal technique requiring that each child be presented each pair of names twice. The formula for the total number of pairings presented in any one group was N (n-1) (n-2).

As well as using the method of reversing the names, the present investigator took special care to assure that the child understood the method and terms involved in making choices.[8] The child, already familiar with the adult, was taken from the group to a special room and given preliminary pairing to assure comprehension of the task involved. Koch used colors or toys as the basis of her preliminary pairings. The present investigator used several pairs of toys, one member of the pair being outstandingly more desirable than the other, for example two red cars, one new and the other broken and missing its wheels; two dolls, one intact and the other missing its head and legs; two aeroplanes, one complete and the other missing one wing and wheels; two doll carriages, one complete and the other without wheels and springs; two small hammers, one ready to use and the other missing the head; and two bells, one complete and one without a clapper. The child was presented with these pairs and asked, "Do you like this one best or that one?" The toys were reversed so that the broken toy was designated once first and once last. Occasionally the child would choose the toy designated as "this" or "that" systematically. In such cases the child was questioned further; did he really like the broken toy and why did he change his choice? The child soon grasped the method involved and would learn to hesitate and look before choosing the toy. It was emphasized that he choose the toy because he liked it best and because it was the most fun to play with. Such pairings were given until the examiner felt sure the child had thorough understanding of the terms and task involved in making choices. The child was then presented with the pairs of names. Friendly encouragement was given as, "Now think," and "Be sure to say the name of the child you like best," with emphasis placed upon the word "best". Often a child spontaneously gave his reason for choosing a name, as "I play with him," "I like him," "I don't like her so I like him best," and so forth. Such remarks were always accepted and the child was encouraged in this kind of verbalization.

To promote good rapport and to avoid fatigue, adjustments were made to fit the individual needs. The number of pairings a subject could take in one sitting varied greatly from subject to subject and from day to day. The child was watched closely and the number of pairings limited accordingly. This was particularly true of Group C, for the larger group entailed many sittings, 10 to 30, and amounted to a total time of from 6 to 12 hours per child. Some children seemed to find satisfaction in the

[8] *The subjects of Group A and C were younger than those used by Koch in her study and it seemed important to be sure these young children understood the method of making a choice.*

task and refused to leave before as many as 60 pairings had been given, instead of the usual 20. Such performance was not encouraged, but, rather than frustrate the child, it was accepted, provided the child seemed really to be interested and was giving true preferences. Not all the children showed this enthusiasm. In such cases the experimenter explained that she had this work to get done and they were helping her. In return for this help she had purchased some toys they might play with. Care was taken to prevent the child from feeling that the pairings were just a means to the goal of playing with toys. Thus some times the toys were played with at the beginning of the interview and the pairings presented when the child felt ready to help the experimenter. Some children seemed to give the pairings more consideration if they sat on the experimenter's lap or held a toy in their hands. In these ways the individual needs of the children were met as far as possible at each sitting.

The teachers were also given the paired comparison choices to determine which children they felt to be the most popular with the other children. They were asked to indicate the child in each pair whom they believed to be the more popular with the children, and were asked not to consider the choice from the adult point of view. Each teacher was given the pairings only once, $\frac{n\ (n-1)}{2}$, for an adult was not expected to favor the first or last name of a pair systematically. To obtain the teacher ranking of child popularity for each group, the two teachers' preferences were combined and then placed in rank order. It was felt the combined scores of two teachers would result in a more valid score than one teacher's estimate. Group B was the exception to this method, for only one teacher's rankings were obtained.

Scoring of Preferences

The child preferences in groups A and B were scored in three ways. One method was the computation of the total number of preferences each child received from all the pairings, both initial and reversed. This was the method used by Koch in her study of popularity (9). The second method used just the initial set of pairings and did not include the reversals. In this method the pairings were so arranged that a particular name appeared as the first member of a pair as many times as it appeared the last member of a pair. As boys of this age seem to favor boys and girls favor girls, the names were presented first as many times as last in all bisex pairs. Only half the total number of pairings were used in this method. The third method of scoring the preferences considered only the pairings in which the child giving the pairings had been consistent, i.e., the same name was chosen as being most liked both times that the pair was presented. These were called consistent pairings, and tended to rule out the pairs in which the child tended to favor the first or last name of a pair.

In order to determine which method of scoring the preferences should be used in the final analysis of data, some investigation was directed toward comparison of these three methods. Intercorrelations, Spearman rank difference, were obtained between these three methods and the teacher ranking of child popularity, for groups A and B. As the half method seemed least reliable as far as groups A and B were concerned, it was not computed for C. For Group C intercorrelations were obtained

between total, consistent and teacher rankings of child popularity.

Table 1 gives the intercorrelations between the three methods of scoring the child preferences and the correlation of each with teachers' estimate of child popularity. In considering the correlations between the total method (the preferences in the initial and reversed order) and teachers' estimate of popularity, the correlation value seemed to decrease

TABLE 1

INTERRELATIONS BETWEEN THE THREE METHODS OF SCORING THE CHILD PREFERENCES AND CORRELATIONS OF EACH WITH TEACHER ESTIMATE OF CHILD POPULARITY

Methods	Group A (N-15)	Group B (N-9)	Group C (N=21)
Total with teacher[9]	.23 ± .17	.54 ± .17	-.07 ± .14
Total with consistent	.90 ± .03*	.80 ± .09*	.91 ± .03*
Teacher with consistent	.33 ± .16	.47 ± .23	.01 ± .14
Teacher (1) with teacher (2)	.79 ± .05*	—	.64 ± .09*
Total with (a) half	.65 ± .11*	.72 ± .12*	—
Total with (b) half	.92 ± .03*	.92 ± .03*	—
Consistent with (a) half	.63 ± .11*	.76 ± .10*	—
Consistent with (b) half	.57 ± .12*	.78 ± .09*	—
Teacher with (a) half	.54 ± .13*	.20 ± .23	—
Teacher with (b) half	.09 ± .18	.73 ± .11*	—
(a) half with (b) half	.02 ± .18	.44 ± .19	—

*Indicates the correlations that are greater than four times their probable error.

with the size of the group. The small group of 9 subjects gave a correlation of .54 ± 17; for the group of 15 subjects, the correlation was .23 ± 17; and the group of 21 subjects yielded a correlation of -.07 ± .14. Koch obtained a correlation of .76 between teacher estimate of child popularity and child estimation.[10] The total score and unanimous (consistent choices) method correlated .90 ± .03, .80 ± .09,[11] and .91 ± .03, for Groups A, B and C respectively. Consistent and teacher estimate of child popularity correlated .33 ± .16, .47 ± .23, and .01 ± .14 for groups A, B and C respectively. When the rankings of the two teachers were separated and compared for Groups A and C the correlations were .79 ± .05 and .64 ± .09 respectively. Group B had only one teacher's ranking, thus there was no inter-teacher correlation for this group.

The total score was divided into the initial and reversal pairings to give the half methods, initial (a) and reversal (b). For further discussion of the half method see page 309. If this method proved satisfactory it would be advantageous, for it required only half the total number of pairings and thus cut the required time of giving the pairings in half. In Group A the half methods (a) and (b) correlated with total .65 ± .11 and .92 ± .03 while they correlated .63 ± .11 and .57 ± .12

[9] *Dr. Koch (9) obtained a correlation of .76 between teacher and total rankings of popularity on a group of 17 children.*

[10] *Dr. Koch suggests that the high correlation obtained between teacher and child estimates of popularity in her study may have been due to the many years of experience and insightfulness of the one teacher giving the rankings for the children in her study.*

[11] *In considering the correlations for group B it is important to remember there were only 9 subjects in that group.*

with consistent. The two halves and teacher estimate of child popularity correlated (a) .54 ± .13 and (b) .09 ± .18. When the two halves were correlated with each other rho was only .02 ± .18. Table 1 gives these same correlations for Group B. On the whole they agree with those of A. As the two half methods showed zero or negligible correlation with each other and were inconsistent in their relationship with the total method, they did not seem worth considering for Group C. When the half methods were thus eliminated, the total and consistent methods remained for further consideration. It was expected that consistent choices were stimulated by real preferences, for this method necessitated the child's choosing of the same name twice, irrespective of position. In comparing the range in scores of these two methods the consistent gave a relatively wider range.[12] Thus for these reasons the consistent method seemed to be slightly more advantageous as a method for scoring the preferences. It was used in final analysis of the data.

Validity of Preferences

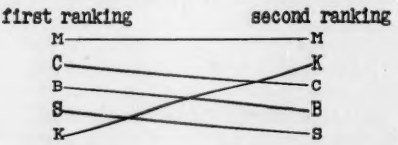
Dr. Koch, in using the total method, found a correlation of .76 between teacher and child estimates of popularity; data on the present groups, A, B and C, did not show this agreement. The correlations for these groups were .23 ± .17 for A, .54 ± .17 for B, and -.07 ± .14 for C. These low correlations may have been prompted by invalid responses on the child's part or the inability of the teachers to estimate child values of popularity. In considering the preferences given by the children of Groups A, B and C as compared with those obtained by Koch in her study, certain similar trends are found. Koch's group gave 46 per cent consistent choices while Groups A, B and C gave 53 per cent, 55 per cent and 41 per cent respectively. Koch found her subjects favored the last name of the pair in 68 per cent of the preferences, while the present groups favored the last name of the pair 56 per cent, 44 per cent and 58 per cent respectively for Groups A, B and C. In the bisex pairs the subjects of Koch's study tended to favor their own sex, boys favored boys in 54 per cent of the bisex pairs and girls favored their own sex in 58 per cent of the bisex pairs. In Groups A, B and C boys favored boys in bisex pairs 52 per cent, 54 per cent and 54 per cent respectively, while girls favored their own sex in 42 per cent, 71 per cent and 57 per cent of the bisex pairs. These discrepancies are perhaps accounted for by the number and popularity of girls in each group. In Group A two popular girls were dropped out of the study on account of prolonged absences; thus the girls left in the study were of less popular ranking. This might account for the girls of this group favoring boys. For in Group B only two girls remained and they happened to be very good friends; thus they tended to choose their own sex in bisex pairings. In Group C where the boy-girl distribution was about equal and only one child was dropped out of the group, the girls favored their own sex in bisex pairings in much the same way as those of Koch's study, as the following tabulation shows:

[12] *Range in scores by total method was A 146-214, B 40-71, C 330-456 and by the consistent method A 31-69, B 7-21, and C 61-117.*

	Group A (N 15)	Group B (N 9)	Group C (N 21)	Koch's Group (N 17)
Last name of the pair favored	56%*	44%	59%	68%
Number of boys in the group	11	7	11	8
Boys favor boys (in bisex pairs)	52%	54%	54%	54%
Girls favor girls (in bisex pairs)	42%	71%	57%	58%
Teachers favor boys (in bisex pairs)	73%		73%	

*The total method of scoring these preferences was used for computing the percentages to enable comparison of the results obtained in Groups A, B and C with those obtained on Koch's group.

To investigate the validity of the child's preferences, 13 subjects of Group C were given 20 pairings again a month and a half after the original testing. These 20 pairings consisted of names of 5 children in all possible combinations and in both the initial and reversed order, n(n-1). In comparing this ranking of these 5 children with the former ranking only one child was misplaced in rank order, but this misplacement was rather a large one, the least popular child moved up to second place. Because of the small number of cases this one displacement reduced the rank difference correlation to .64 ± .20. The change in rank order from the first to the second ranking according to total method is shown below. The number was too small to attempt the consistent ranking.



When Group D, the 10 follow-up cases of Group A, were retested 9 months later, the pairings obtained at this later date correlated with the first .50 ± .17, in spite of the fact that these 10 children were not often together in the preschool group. The routine of the preschool necessitated a subdivision of the group for nap time, which placed 5 of Group D in one sleep group and the other 5 in the other. Although the 10 children did see each other from time to time, they did not play together each day.

Summarizing the validity and reliability of the children's choices, there seemed to be evidence that the popularity scores were not due to chance alone, but that real preferences on the children's part to some extent prompted their choices. The discrepancy between child and teacher rankings of child popularity did not rest entirely upon inadequate responses on the child's part. Might it not in part have been due to teacher estimate of child popularity?

The 5 adults involved in the estimating of the child popularity were

well trained preschool teachers, and had worked in the preschool some months prior to giving their preferences for child popularity. In combining the paired comparison preferences of two teachers, it was felt that a more valid estimate was obtained than by taking just one teacher's ranking. The tabulation on page 312 shows that teachers favored boys in 73 per cent of the bisex pairs, in both groups, when the teachers gave their estimates of child popularity by the paired comparison method. The following tabulation shows that the children seem to give approximately an equal mean number of preferences to boys as to girls.

	Group A	Group B	Group C
Number of boys	11	7	11
Number of girls	4	2	10
Consistent Preferences			
Mean number boys received	50.90	16.27	78.40
Mean number girls received	43.25	16.50	85.60
Teacher Preferences			
Mean number boys received	15.72	*	26.18
Mean number girls received	9.25		16.10

*Teacher preferences were not obtained for Group B.

If anything, girls receive a few more choices than boys, as Koch found in her study. The popular girls in Group A were dropped out on account of prolonged absences; thus it was expected that boys might receive a few more preferences than girls in that group.

Whether the teachers really showed sex preferences or not is difficult to say. The teachers may have been led to think boys more popular than girls by the fact that boys were higher in social participation than girls. Teacher estimate of child popularity and intensity of social participation correlate .81 ± .06 and .63 ± .09 for Groups A and C respectively (See Table 2), while children's ranking of popularity, consistent choices, and social participation correlate .48 ± .14 and .07 ± .14 respectively (See Table 4). It is possible that this difference in social participation accounted for the sex difference shown by the teachers in their estimate of popularity.

When the two teacher estimates of child popularity were compared for Groups A and B the correlations were .79 ± .05 and .64 ± .09 respectively. Although the teachers scored their preferences independently, they seemed to show considerable agreement.

It seems most probable that the discrepancy between child and teacher estimates of child popularity was not due to unreliable measures, but that the two measures were based upon different criteria. According to Hsia (6) who studied popularity of grade school children, one should not expect teacher and child ratings of popularity to resemble each other very closely, for teacher criteria for judging popularity are bound to differ from those of the children. In such a case it seemed advisable to use both measures of popularity for final analysis of data.

Observation Blank of Social Interactions

The purpose of the observations was to obtain data pertaining to social behavior of each child while in the preschool situation. The items observed were: 1) intensity of social participation, a weighted score derived from the degree to which the child was a group member, i.e., each half minute of the 80 observed minutes that was spent in isolated play was given a score of zero, each half minute spent in peripheral play was given a score of one-half, and each half minute spent in active or leadership play was given a score of one; 2) the number of times, during the observations, that the child gave and received directing, compliant, resistant and communicative actions; 3) the quality of each interaction (both given and received) whether friendly, matter-of-fact, or hostile; and 4) the children played with. For further definitions of these terms, for methods of recording, and for the observation blank see page 88 of reference (10).

The observations were taken in 10-minute records for each child. It was felt that any smaller units of observations would prevent the observer from obtaining information pertaining to intensity of social participation. Eight records were found to be a reliable sampling as compared to 16 such records. Reliability of the sampling and observers, as obtained between two independent observers for 15 observations, is given in the following tabulation. The reliability of the two observers was obtained from the formula $\frac{2 \times \text{number of agreements}}{\text{total number of items recorded}}$. The reliability of the sampling was obtained by rank difference correlations between 8 and 16 observations for 13 subjects.

Item	Reliability of Observer	Item	Reliability of Sampling
Isolated play	.90	Intensity of social participation	.91
Peripheral play	.81		
Active-leader play	.82		
Directed others	.75	Directed others	.73
Others directed	.99	Others directed	.87
Complied to others	.75	Compliant behavior	.88
Was complied to	.96		
Refused others	.77	Refusing behavior	.73
Was refused	.80		
Friendly behavior	.94	Friendly behavior	.80
Hostile behavior	.81	Hostile behavior	.86
Child-child conversation	.80	Child conversation	.94
Teacher-child conversation	.89	Teacher-child conversation	.98

The observations were made when the children were engaged in free play, thus allowing optimum opportunities for social interactions. Care was taken not to obtain more than one record for a child on any one day. Absences and the school testing program prevented any systematic order of observation of the children, but emphasis was placed upon beginning the observations on all members of a group the same week. Half the observations were taken indoors and half outdoors, for each child.

Merrill-Palmer Rating Scales

Merrill-Palmer rating scales (4) and (13) were used for ratings of attractiveness of personality, co-operation in routines, and physical attractiveness. These scales are so constructed that the rater may check a few items or quite a number. The child's score is the mean value of the various items checked, as obtained from two teachers' ratings.

Constructive-Destructive Test

The test of constructive-destructive tendencies was developed by Ackerman (1) and (2). This test confronted the child with a realistic play situation in which the choice of action was limited to constructive or destructive behavior, in 10 different play units. Constructiveness is operationally defined for the purpose of this test as building up of toys, attempting to place pieces together to form a structure, as for example, placing the head on the doll and placing a car on its wheels. Destructiveness is operationally defined as breaking down of play units, taking the set of toys apart or trying to break the pieces. Any investigating behavior that removes part of the toy is classed as destructiveness. The score varies from a minus ten to a plus ten, and is obtained for the initial approach of the child to each play unit and by the final appearance of the toy units when the child is finished. Thus two scores were obtained, one for the initial behavior and one for the final constructiveness. Detailed description of the play units and validity and reliability is given in reference (10).

Gross Motor Tests

An estimate of the child's gross motor ability was obtained by using some of the McCaskill motor items (11), namely, ascending and descending steps and ladders. The child was scored on his performance in two of the three trials on each item. For further description of these items see (10).

Paper Cutting Test

The objectives in choosing a fine motor test were to find items that the child would enjoy at this age level, that would not show sex differences, that would discriminate between abilities of three year olds, and that would not correlate highly with the McCaskill (11) ascending and descending stairs and ladder items. Jakway's (7) paper

cutting test seemed to answer these requirements. It required the child to cut along a line 1/4 inch wide with blunt scissors. The test had three degrees of difficulty, a straight line, an S shaped object, and a star.

Nervous Habits

The observing and recording of oral nervous habits[13] were patterned after Carr's (3) revision of Olson's (12) methods. Each child was observed for one minute a day for 30 days. The presence or absence of any oral nervous habit within that time-sampling was recorded. The score was the number of minutes in which oral nervous habits occurred. Reliability of the sampling and observers are given on page 103 of reference (10).

Intelligence Test

The subjects were given the revised Stanford Binet test of intelligence, Form L. The examiners were experienced testers of the Iowa Child Welfare Research Station of the University of Iowa. The child, already familiar with the tester, was examined on the day of the month corresponding as closely as possible to the day of the month upon which his birthday fell. Mental ages were computed for the midpoint of the period during which the data were being collected.

Failure Test

The subject's reaction when confronted with failure was measured by the Keister (8) puzzle box test of failure. The task of the test was to replace some objects into the box and close the lid. This seemed like a simple task to the child, but due to the shape of the figures and the size of the box it presented a difficult problem even for an adult, and thus the child experienced failure. The observer recorded the child's overt behavior by checking each half minute for the occurrence of the following behavior items: attempts to solve alone, asks another to solve, asks help, destructive behavior, stops trying, succeeds, and rationalizes. Further description of overt behavior was recorded each half minute under the headings: manifestations of anger, whines, cries, no manifestations of anger, sulks and yells.

Since Keister (8) and Goodman (5) classed their subjects into groups rather than giving them raw scores, the scoring method was changed for the present study. The test seemed to obtain results on two types of failure behavior, constructiveness and emotional response. These two types of behavior were scored separately for Groups A, B and C. One score was obtained by subtracting the number of half-minutes the child spent in nonconstructive behavior from the number of half-minutes the child spent in constructive efforts to solve the puzzle. The second score involved the emotional manifestations of behavior, and was obtained by subtracting the number of half-minutes the child spent in showing no

[13] Olson operationally defines oral nervous habits as thumbsucking, picking of teeth, finger or hand sucking, biting fingernails and sticking out the tongue.

manifestations of emotional behavior from the number of half-minutes spent in manifestations of anger, crying, and so forth. The score on each type thus ranged from a -30 to a +30.

RESULTS

As child and teacher estimates of child popularity seemed to be based upon different criteria, it was felt that analysis of the data in terms of both popularity rankings should yield different results. The data were dealt with by two methods to show the relationship between popularity and the other measures: 1) rank difference correlations were computed between the two criteria of popularity and the other measures; 2) Fisher "t's" were determined between the means of the upper, popular third and lower, unpopular third of each group. These comparisons were computed for each of the three groups separately, and for the combined groups, according to both measures of popularity.

In Group D rank difference correlations were used to show the relationship between initial and final measures, taken at an interval of 9 months, and also to relate change in popularity to change in other measures.

Relationship of Teacher Ranking of Child Popularity to Other Measures

Rank Difference Correlation Method

Table 2 gives the correlations between teachers' estimate of child popularity and the other measures. The following measures correlated positively with teachers' estimate of child popularity in all three groups: 1) intensity of social participation; 2) the number of times the child directed others; 3) the frequency with which the child was directed by others; and 4) the number of times the child was compliant to others. The correlations between intensity of social participation and popularity were .81 $\pm$ .06,[14] .33 $\pm$ .21, and .63 $\pm$ .09 for Groups A, B and C respectively. Group B's correlation was low, but there were only 9 subjects in that group and only one teacher's ranking of popularity was obtained. These correlations indicate that the child who showed a relatively large amount of active participation in a group was likely to be popular as judged by teachers. The correlations between popularity and directing others were not high, but show similar trends. They were .31 $\pm$ .16, .23 $\pm$ .22, and .68 $\pm$ .08 for Groups A, B, and C respectively. Between receiving of directions and popularity the correlations were .40 $\pm$ .15, .55 $\pm$ .17, and .53 $\pm$ .10 respectively for Groups A, B and C. Compliant behavior, as obtained from observations of the child at play, seemed to be more related to popularity than were teacher ratings of child co-operation in routines. The former, compliant behavior, correlated with popularity .52 $\pm$ .13, .60 $\pm$ .15 and .61 $\pm$ .09 for the three groups respectively. The latter, co-operation in routines, correlated with popularity .42 $\pm$ .15, .78 $\pm$ .09 and -.24 $\pm$ .14 for the three Groups A, B and C.

[14] *Probable error is given for all correlations.*

TABLE 2

RELATIONSHIP BETWEEN TEACHERS' ESTIMATE OF CHILD POPULARITY AND THE OTHER MEASURES

Measure	Group A N = 15	Group B N = 9	Group C N = 21
Constructive-destructive test			
Initial approach	-.08 ± .18	.03 ± .24	.25 ± .14
Final score	.22 ± .17	.15 ± .23	.01 ± .15
Merrill-Palmer ratings			
Attractiveness of personality	..23 ± .17	.48 ± .18	.06 ± .14
Cooperation in routines	.42 ± .16	.78 ± .08	-.24 ± .14
Physical attractiveness	-.33 ± .16	.93 ± .03	-.27 ± .13
Motor Tests			
Gross	.12 ± .18	.06 ± .24	.38 ± .12
Paper cutting	.61 ± .11	.25 ± .22	.23 ± .14
Failure Test			
Constructiveness	-.15 ± .17	-.27 ± .22	-.07 ± .14
Emotional manifestations	-.23 ± .17	-.30 ± .21	-.18 ± .14
Preschool attendance	.28 ± .16	.39 ± .20	-.10 ± .14
Chronological age	.28 ± .16	.20 ± .23	.08 ± .14
Mental age	.03 ± .18	.50 ± .17	.32 ± .13
Intelligence Quotient	.04 ± .18	.43 ± .19	.25 ± .14
Nervous habits	-.08 ± .18	-.08 ± .23	.28 ± .13
Communication			
Child to child	.28 ± .17	.28 ± .23	.76 ± .03
Child to teacher	-.17 ± .18	-.20 ± .23	-.11 ± .03
Intensity of social participation	.81 ± .06	.33 ± .21	.63 ± .09
Social behavior			
Directed others	.31 ± .16	.23 ± .22	.68 ± .08
Was directed by others	.40 ± .15	.55 ± .17	.53 ± .10
Complied to others	.52 ± .13	.60 ± .15	.61 ± .09
Was complied to by others	.18 ± .18	.20 ± .23	.63 ± .09
Refused others	-.18 ± .18	.76 ± .10	.06 ± .14
Was refused by others	.48 ± .14	-.01 ± .24	.38 ± .12
Received friendly behavior	.43 ± .15	—	.04 ± .14
Gave friendly behavior	.12 ± .18	.41 ± .20	-.21 ± .14
Received hostile behavior	-.26 ± .17	-.04 ± .24	.38 ± .12
Gave hostile behavior	.11 ± .18	-.08 ± .23	.59 ± .09

The following measures seemed to be unrelated or very slightly related to popularity as rated by teachers; 1) constructive-destructive behavior as measured by Ackerman's test; 2) attractiveness of personality as measured by teacher ratings; 3) gross motor ability as measured by ascending and descending stairs and ladders; 4) the number of oral nervous habits; 5) intelligence quotient and mental age; 6) the number of days the child attended preschool during the semester in which the

data were gathered; 7) amount of communication with the teacher; 8) the amount the child refused to comply to others; 9) the amount of friendly and hostile social behavior the child received; 10) the amount of friendly behavior the child gave to others; 11) reaction when faced with failure; and 12) chronological age. The small age range of each group may have been somewhat responsible for the low correlations between age and popularity.

Some group differences appeared in the relationship of certain of these measures to teachers' ranking of popularity. In Group A popularity and paper cutting ability correlated .61 $\pm$ .11, while in Group B and C the correlations were low between these two measures. In Group B popularity correlated .93 $\pm$ .03 with physical attractiveness, .76 $\pm$ .10 with the number of times the child refused to comply in social situations, and with co-operation in routine .78 $\pm$ .08, while the correlations were low between these measures for Groups A and C. Group C was the only group for which the correlations showed a substantial relationship between the following items and popularity: 1) the amount of communication with other children correlated .76 $\pm$ .03; 2) the number of times the child was complied to by others correlated .63 $\pm$ .09; 3) the amount of hostile behavior the child gave to others correlated .59 $\pm$ .09; and 4) the number of times the child directed others correlated .68 $\pm$ .08.

Comparison of Means of Popular and Unpopular Subjects

Table 3 gives the means of the upper, middle and lower thirds of each experimental group on the basis of teacher ranking of child popularity. The level[15] at which we may be confident that the difference between the upper and lower thirds was significant is given just below the mean of the middle third, M_2. In Group A there were significant differences between the popular and unpopular groups on the following measures; 1) fine motor co-ordination, at the 3 per cent level; 3) intensity of social participation at the 1 per cent level; 3) and receiving friendly behavior at the 2 per cent level of confidence. The four measures on which the popular third were significantly higher in Group B are: 1) personality attractiveness, at the 1 per cent level; 2) co-operation in routine, at the 2 per cent level; 3) physical attractiveness, at the 1 per cent level; and 4) the number of times the child resisted others, at the 5 per cent level. This latter finding means that the popular child resisted others a greater number of times than did the unpopular children. Significant at the 1 per cent level, the following items differentiated the popular and unpopular members of Group C: 1) the amount of communication with other children; 2) the amount the child was directed by others; and 3) the amount the child complied to others and was complied to by them. Also in favor of the popular children of Group C, and significant at the 5 per cent level, were: 1) gross motor ability; 2) the frequency with which the child was friendly in social relations with other children. Intensity of social participation was significant at the 4 per cent level in Group C, and direction of others at the 3 per cent level.

When the upper thirds of the three groups were combined and compared to the combined lower thirds, there were significant differences at the

[15] *The level of confidence of "t" represents the chances in one hundred that one would get an obtained value of "t" of this from the same population.*

TABLE 3

COMPARISON OF THE MEANS OF THE POPULAR, UNPOPULAR AND MIDDLE THIRDS OF THE THREE EXPERIMENTAL AND THE COMBINED GROUP, ACCORDING TO TEACHERS' ESTIMATE OF CHILD POPULARITY

Measure	Group A			Group B			Group C			Combined Group A-B-C		
	M1	M2	M3	M1	M2	M3	M1	M2	M3	M1	M2	M3
Constructive-destructive test Initial approach	- 2.0	.57	- .80	.0	- 0.8	1.0	.0	- 2.4	- 2.71	- .57	- .75	- 1.56
Final score	4.0	2.0	- 2.5	7.67	.67	2.0	1.86	2.7	2.71	3.71	1.90	1.07
Merrill-Palmer Ratings Attractiveness of personality	61.0	65.0	48.22	71.67	41.7 1%	51.33	75.43	65.7	73.43	69.5	59.52	61.49
Cooperation in routines	76.0	75.42 10%	52.5	71.67	64.0 5%	57.33	66.29	58.29	77.14	70.71	64.70	61.50
Physical attractiveness	66.0	70.43 8%	73.8	64.33	64.67 1%	47.67	74.86	72.7	79.71	74.86	70.55	71.14
Motor Tests Gross	23.5	24.9	22.75	10.0	8.7	11.33	28.71	19.7 5%	20.14	22.15	19.77 7%	18.94
Paper cutting	5.0	4.0 5%	5.0	11.67	8.67	8.7	4.0	4.0	5.1	5.93	4.92	4.29
Failure Test Constructiveness	5.0	2.86	5.0	- 7.6	-11.33	- 1.33	4.14	1.86	5.86	2.92	- .327	3.50
Emotional manifestations	- 0.5	4.0	4.0	-12.3	-15.33	- 9.0	7.9	.29	14.3	1.15	- 1.51	6.36
Days of Preschool Attendance	63.5	63.50	58.5	56.67	54.33	50.67	62.0	62.29	62.2	61.71	61.34	58.64
Chronological Age	49.25	47.38	50.5	52.0	55.3	50.33	45.43	48.7	45.43	49.07	47.11	49.07
Mental age	64.42	65.14	65.55	79.98	77.81	68.97	61.19	53.75	54.86	63.44	62.28	60.84
Intelligence Quotient	130.8	137.77	129.8	129.0	140.7	111.0	140.9	125.0	128.5	135.4	132.20	124.00
Nervous Habits	11.25	11.42	14.0	13.67	13.67	15.67	7.14	6.57	9.29	9.70	10.39	12.05
Communication Child to child	85.0	69.42	62.0	120.0	72.9	108.5	133.14	93.86 1%	45.42	116.85	80.03 1%	82.82
Child to teacher	28.8	28.42	49.5	71.1	58.4	89.3	48.5	69.42	53.1	47.60	50.24	59.57
Intensity of Social Participation	120.6	97.6 1%	64.1	101.9	64.0	91.4	108.6	86.9 4%	72.9	110.57	89.79 1%	74.97
Social Behavior Directed others	23.18	14.27	14.9	24.0	10.0	18.9	19.2	16.29 2%	5.7	21.36	15.79 1%	10.70
Was directed by others	17.5	10.84	9.0	24.0	10.2	12.0	14.9	9.43 1%	6.7	17.57	9.6 1%	8.51
Complied to others	9.4	6.57	4.5	12.0	2.0	4.4	9.9	5.14 1%	3.7	10.19	5.46 1%	4.02
Was complied to by others	9.67	6.57	6.2	21.5	0.0	9.2	8.57	5.57 1%	2.4	10.61	5.01 2%	5.10
Refused others	1.75	.86	3.8	3.6	0.9 5%	0.9	2.1	3.0	1.7	2.84	1.74	2.12
Was refused by others	4.8	2.71	2.5	1.3	0.0	1.8	2.9	3.3 9%	1.3	3.10	2.60 3%	1.67
Received friendly behavior	7.8	1.8 2%	2.8	1.3	0.0	0.4	0.7	1.0	1.0	2.85	1.11	1.38
Gave friendly behavior	4.5	3.5	3.8	1.8	0.0	0.9	0.0	1.3 3%	1.3	1.67	1.278	1.90
Received hostile behavior	11.5	9.0	9.0	3.1	2.3	4.0	3.1	5.6 4%	0.7	5.5	6.748	3.78
Gave hostile behavior	10.5	7.7	19.8	2.2	5.3	2.1	8.0	7.6 5%	0.3	6.46	7.82	6.43

1 per cent level, in favor of the upper third for: 1) intensity of social participation; 2) the amount of directing behavior toward others; 3) the amount of directing received; 4) the amount of communication with other children; and 5) the amount of compliance to others. Also in favor of the popular children and significant at or below the 5 per cent level are the following: 1) the amount others complied to the observed child; 2) the amount others resisted or refused to comply to the given subject.

In some cases the middle mean of the combined group was greater or smaller than either of the other means, but the difference was not significant. However, when the difference between the upper and lower thirds was significant, the middle mean was always smaller than the upper mean and greater than the lower mean.

Summary of the Relationship Between Teachers' Estimate of Popularity and the other Measures

In summary, the relationships between the 27 measures and teachers' estimates of child popularity fall into three main classifications; (a) measures clearly related to the teachers' estimates, (b) measures possibly

related to teachers' estimates and (c) measures not related to teachers' estimates in this study. To be considered as related to popularity the measures met the following qualifications: 1) it correlated positively in all three groups and in either Group A or Group C[16] the correlation exceeded four times the probable error; 2) the upper popular third was significantly superior to the lower (at the 5 per cent to 1 per cent level) in Group A or C; and 3) when the combined upper thirds of all three groups were compared to the combined lower thirds, the difference was significant at the 5 per cent to 1 per cent level. All measures that did not show as much relationship as this with teacher estimate of popularity, but showed some relationship, perhaps in one group and not in the others, were classed under (b). The measures that showed no agreement, either positive or negative with popularity in any one of the groups, were classed as (c), unrelated to teachers' estimates of popularity.

The measures that met the requirements of classification (a), related to teachers' estimate of popularity, were: 1) intensity of social participation; 2) the amount the child directed others; 3) the amount the subject was directed by others; and 4) the amount the observed child complied to others; and 5) the amount of child-child communication; and 6) the amount the child was complied to by others.

The measures that seemed to be related to teacher estimate of popularity in one or two of the groups, but not in all three, classification (b), were: 1) co-operation in routines; 2) physical attractiveness; 3) attractiveness of personality; 4) fine and gross motor ability; 5) the amount the child was refused by others; 6) refusing to comply in social situation; 7) mental age; 8) the amount of friendly interactions given and received; and 9) the number of hostile social interactions given and received.

The following measures seemed unrelated to teacher estimate of child popularity, and thus fell into classification (c): 1) initial and final behavior on the constructive-destructive test; 2) constructive and emotional behavior in the face of failure; 3) intelligence quotient; 4) chronological age; 5) the number of days attended the preschool during the semester the data were gathered; 6) the number of oral nervous habits manifested; and 7) the amount of communication between teacher and child.

Relationship of Other Measures to Child Ranking of Popularity

This popularity ranking was based on preferences made by preschool children of their associates. The child values underlying these preferences were most likely resultants of many pleasant and unpleasant social experiences. The teacher, on the other hand, probably searched for some outstanding behavioral characteristics upon which to base her judgments. From an adult's point of view it seemed plausible that the child who was the leader or socially an active group member was popular

[16] *Group B had only nine cases and for this reason the results are not given as much weight as those of Groups A and C for which the findings should be more reliable.*

with the children, and an aggressive child would appear more popular than the quiet retiring type of child. With the children, however, such a characteristic as social participation alone was not sufficient to render the child popular. It seemed that the child chosen as popular by his associates differed in several respects from the child chosen as popular by the teacher.

Relationships between Child Popularity and the Other Measures

Table 4 presents the rank difference correlations that were obtained between child estimates of popularity and the other measures. The three groups seem to substantiate Koch's finding that high scores on routine co-operation were more characteristic of popular than of unpopular children. The correlations between routine co-operation and popularity for Groups A, B and C were .44 ± .15, .65 ± .14 and .24 ± .14 respectively. These correlations were not as high as the one Koch obtained for her group, .92 ± .02, but were in the same general direction. The correlations between popularity and constructiveness in the face of failure were consistently negative, but only Group A's was significant,[17] -.53 ± .13; Group B's was -.30 ± .22; and Group C's was -.04 ± .14. These correlations indicate that the popular child tended to be destructive in the face of failure, as compared with the unpopular child.

The following measures seemed to be unrelated to popularity: 1) final score on the constructive-destructive test; 2) paper cutting ability; 3) intelligence quotient; 4) the number of days attending preschool the semester the data were gathered; 5) chronological age; 6) emotional reaction to failure; 7) the amount of directing given and received; 8) amount of communication between children; 9) the amount of complying and being complied to; 10) the amount of refusing to comply; 11) the number of friendly interactions given and received; and 12) the number of hostile interactions given and received.

When the results on the three groups were compared some differences were found to exist. In Group A constructiveness, as indicated by initial approach score, correlated with popularity -.69 ± .10, i.e., the popular children were destructive as compared to the unpopular children. However in Groups B and C the correlations were low and positive. In Group B popularity and physical attractiveness correlated .52 ± .17, but in Group C the correlation was lower and in Group A it was zero. The relationship between gross motor ability and popularity seemed to show inconsistent results. In Group A the correlation was .58 ± .12, while B's was -.54 ± .17, and C's was .09 ± .14. The relationship between the amount of teacher-child interactions and popularity also varied for the three groups. Group A showed no relationship, B showed a negative relationship of -.50 ± .18, and the data for Group C indicated a positive correlation of .54 ± .09. Group A was the only group to show a relationship between popularity and attractiveness of personality, and between popularity and intensity of social participation; the correlations were .45 ± .15 and .48 ± .14 respectively. Group B

[17] *The term significant is used here to indicate correlations that are greater than four times their probable error.*

TABLE 4

RELATIONSHIP BETWEEN CHILD POPULARITY AND THE OTHER MEASURES

Measure	Group A N = 15	Group B N = 9	Group C N = 21
Constructive-destructive test			
Initial approach	-.69 ± .10	.40 ± .20	.18 ± .14
Final score	-.26 ± .17	.16 ± .23	.26 ± .13
Merrill-Palmer Ratings			
Attractiveness of personality	.45 ± .15	-.12 ± .23	.28 ± .13
Cooperation in routines	.44 ± .15	.65 ± .14	.24 ± .14
Physical attractiveness	.00 ± .18	.52 ± .17	.31 ± .13
Motor tests			
Gross	.58 ± .12	-.54 ± .17	.09 ± .14
Paper cutting	.36 ± .16	.19 ± .23	.20 ± .14
Failure test			
Constructiveness	-.53 ± .13	-.30 ± .22	-.04 ± .14
Emotional manifestations	-.36 ± .17	-.21 ± .23	.13 ± .14
Days of preschool attendance	-.30 ± .17	.10 ± .23	.22 ± .14
Chronological age	.06 ± .18	-.33 ± .21	.18 ± .14
Mental age	.34 ± .15	.28 ± .21	.19 ± .15
Intelligence Quotient	.31 ± .17	.38 ± .20	.07 ± .14
Nervous habits	.12 ± .18	.52 ± .17	.17 ± .14
Communication			
Child to child	.07 ± .18	.27 ± .22	-.11 ± .14
Child to teacher	-.08 ± .18	-.50 ± .18	.54 ± .09
Intensity of social participation	.48 ± .14	.08 ± .24	.07 ± .14
Social behavior			
Directed others	.31 ± .17	-.28 ± .22	.01 ± .14
Was directed by others	.01 ± .18	-.17 ± .23	-.15 ± .14
Complied to others	-.01 ± .18	-.20 ± .23	.13 ± .14
Was complied to by others	.25 ± .17	-.33 ± .21	.08 ± .14
Refused others	.10 ± .18	.20 ± .23	-.09 ± .14
Was refused by others	.39 ± .15	-.55 ± .17	.11 ± .14
Received friendly behavior	-.04 ± .18	.43 ± .19	.09 ± .14
Gave friendly behavior	.12 ± .18	-.09 ± .23	-.08 ± .14
Received hostile behavior	.17 ± .18	-.44 ± .19	.00 ± .14
Gave hostile behavior	-.14 ± .18	.00 ± .24	-.03 ± .14

was the only group in which there was an indication of a relationship between popularity and mental age; popularity and the amount others refused to comply; and popularity and manifestation of nervous habits. The correlation was .53 ± .17 for mental age, .52 ± .17 for attractiveness of personality, -.55 ± .17 for others refusing to comply, and .52 ± .17 for nervous habits.

Significance of the Difference Between the Means of the Upper, Popular and Lower Unpopular Thirds on the 27 Measures

Table 5 gives the means of the measures for the upper, middle and lower thirds, according to child ranking of popularity. When the difference between the upper and lower thirds was significant at the 10 per cent to 1 per cent level, the "t" has been indicated below the middle mean in Table 5. From this table it was apparent that there was no one measure that showed a significant difference at or below the 10 per cent level for all three groups treated separately, but when the upper thirds of the three groups were combined and compared to the combined lower thirds there was a difference significant at the 10 per cent level for co-operation in routines. Thus there was a slight indication that the popular children were rated as being more co-operative in routines than were the unpopular.

In considering the groups individually there were significant differences at the 1 per cent and 5 per cent levels in Group A for initial

TABLE 5

COMPARISON OF THE MEANS OF THE POPULAR, UNPOPULAR AND MIDDLE THIRDS OF THE THREE EXPERIMENTAL AND THE COMBINED GROUPS, ACCORDING TO CHILD POPULARITY

Measure	Group A			Group B			Group C			Combined Group A-B-C		
	M1	M2	M3	M1	M2	M3	M1	M2	M3	M1	M2	M3
Constructive-destructive test Initial approach	- 4.0	- 2.4 1%	5.2	3.7	- 3.7	- 1.33	- 2.3	- 2.0	- 0.9	- 1.67	- 2.25	1.06
Final score	- 0.4	- 2.8 5%	7.2	6.3	5.0	3.3	3.8	0.0	- 1.1	2.80	.06	2.53
Merrill-Palmer Ratings Attractiveness of personality	64.2	55.4	56.0	50.55	48.3	68.0	72.3	68.0	70.3	65.20	59.46	65.10
Cooperation in routines	74.0	68.4	63.4	72.7	55.0 1%	47.33	69.7	64.8	65.5	71.75	63.33 10%	61.11
Physical attractiveness	70.8	67.8	72.0	74.0	64.0	59.75	78.4	70.4	72.4	73.90	71.90	69.5
Motor Tests Gross	27.0	23.2 5%	21.6	6.7	12.0 1%	11.3	24.6	17.9	23.3	21.75	18.45	20.3
Paper cutting	4.8	3.6	3.8	9.7	10.7	9.0	4.0	3.7	3.4	5.30	5.16	4.86
Failure Test Constructiveness	- 5.5	0.0 5%	12.2	- 4.7	-11.3	- 4.3	1.7	2.87	2.3	1.93	- .86	4.97
Emotional manifestations	-11.0	9.2	7.2	-11.3	-14.0	-11.3	8.1	10.6	6.0	- 2.1	5.16	2.93
Days of preschool attendance	59.8	60.8	65.8	52.3	53.7	53.7	63.0	64.7	59.1	59.80	61.59	60.3
Chronological age	49.6	48.0	48.4	58.0	54.3	61.3	43.1	43.0	42.0	48.90	48.98	48.0
Mental age	66.58	67.1	61.46	70.65	65.15	76.98	54.83	55.9	54.6	64.09	61.96	61.77
Intelligence Quotient	134.2	139.8	127.0	137.3	120.0	127.7	129.7	130.0	130.4	132.70	131.25	129.7
Nervous habits	11.6	11.8	13.0	12.3	14.0	15.7	7.1	7.3	8.7	9.67	10.05	11.74
Communication Child to child	75.0	74.4	67.4	101.1	82.6	112.0	91.3	89.4	111.7	87.1	73.71	96.99
Child to teacher	82.0	86.8	82.4	82.0	100.9	89.3	68.3	85.7 2%	87.4	49.0	65.09	86.1
Intensity of social participation	109.9	78.9	88.7	75.45	92.9 10%	86.4	94.3	86.3	89.9	95.8	85.11	87.1
Social behavior Directed others	20.5	18.4	15.4	15.8	12.4	22.7	15.4	12.1	13.7	17.13	14.34	15.40
Was directed by others	10.06	14.2	11.2	14.2	13.3	18.7	11.87	6.85	14.4	11.80	10.39	14.20
Complied to others	5.5	6.8	6.2	4.2	3.5	8.1	8.4	4.4	7.9	5.75	5.04	7.87
Was complied to by others	8.7	8.0	5.0	4.3	11.1	10.7	6.9	4.9	4.9	6.98	7.14	6.06
Refused others	2.9	0.8	2.0	1.3	2.2	1.0	2.3	1.4	3.1	2.40	1.44	2.20
Was refused by others	4.1	3.4	2.2	0.0	1.3 4%	1.7	3.1	2.1	3.14	2.90	2.40	2.53
Received friendly behavior	4.0	2.6	4.0	0.7	0.9	0.0	0.8	1.4	0.9	1.85	1.69	1.90
Gave friendly behavior	4.0	3.2	4.0	0.3	0.9	1.0	0.9	0.4	1.8	1.80	1.44	2.8
Received hostile behavior	9.4	9.2	10.4	2.3	2.7	2.7	3.3	2.4	3.7	5.30	4.71	5.80
Gave hostile behavior	9.4	16.8	8.8	4.4	3.5	2.3	4.8	3.0	5.87	6.15	7.68	5.80

and final constructive-destructive scores, in favor of higher constructiveness in the unpopular group. The popular children of Group A were superior to the unpopular in gross motor ability, significant at the 4 per cent level. Also significant at the 4 per cent level is the reaction of Group A's members to failure. Here again the unpopular children were more constructive. Group B's popular children were refused by others in social situation, significant at the 4 per cent level and co-operated in routines significant at the 3 per cent level. In this same group gross motor ability and teacher communication were significant at the 1 per cent and 10 per cent levels respectively. In Group C the popular children had more social contacts with the teacher than the unpopular third, significant at the 2 per cent level.

Summary of the Relationship Between Other Measures and Child Popularity

Summarizing the relationship between popularity and the other measures, routine co-operation as measured by teacher ratings seemed to be the one measure that was positively related to popularity in all three groups and for which the correlation was four times its probable error in at least one of the groups. The following measures seemed unrelated to popularity as far as these three groups were concerned: 1) the number of days of attendance during the semester the data were gathered; 2) chronological age; 3) emotional reactions to failure; 4) the amount of giving and receiving of directing actions; 5) the amount of communication with other children; 6) the giving and receiving of compliant actions in social situation; 7) the number of social interactions given and received of both friendly and hostile type; and 8) the number of nervous habits manifested.

The following items showed inconsistency in their relationship to popularity, i.e., in one group the relationship might have been positive, in another negative, and in the third there might have been no apparent relationship: 1) constructive-destructive behavior in the test situation; 2) attractiveness of personality; 3) physical attractiveness; 4) gross motor ability; 5) the amount of communication with teacher; and 6) the amount of refusing and of being refused by others. The following measures showed consistent low positive correlations with popularity: 1) paper cutting ability; 2) intelligence quotient; 3) mental age; and 4) intensity of social participation. The constructive approach to failure tended to be negatively related to popularity.

From Table 5 it may be noticed that the difference between the highest and lowest thirds of the combined groups, although not significant, tended to show that the popular children were more co-operative in routines, had a higher mental age, and were of greater physical attraction than the unpopular children; while the middle third seemed to be superior to the others in showing better emotional reaction to failure, but tended to be inferior in chronological age, in constructiveness, and in gross motor ability.

Results of the Measures on Group D

Ten members of Group A, 3 girls and 7 boys, remained in the preschool the following year and were retested after an interval of nine

months. The scores, taken when the children were members of Group A, were recalculated to give inter-rankings for the 10 subjects and compared by rank difference correlations to the scores obtained at the later date. (See Table 6.) The change in popularity, from initial measures, was related to change in scores of the other measures by rank difference correlations.

Unfortunately the number of cases is very limited, but as far as these 10 cases were concerned some of the measures seem to yield consistent results over the 9 months' period. The final score on the constructive-destructive test correlated .76 ± .09 between the initial and final measurement, but the score pertaining to first approach in this test only correlated .38 ± .19 for the two testings. Due to a change in teaching staff, the two sets of Merrill-Palmer ratings were not obtained from the same teachers. The only one of these ratings that showed

TABLE 6

COMPARISON OF TEST AND RETEST DATA AFTER A NINE MONTHS' INTERVAL (N = 10)

Measure	Correlation	Measure	Correlation
Constructive-destructive test		*Communication*	
Initial	.38 ± .19	Child-child	.72 ± .11*
Final	.76 ± .09*	Child-teacher	.39 ± .19
Ratings		Intensity of social participation	.50 ± .17
Personality attractiveness	.14 ± .22	Social behavior	
Cooperation in routine	.39 ± .19	Directed others	.78 ± .09*
Physical attractiveness	.75 ± .10*	Was directed	.65 ± .13*
Motor tests		Complied to others	.35 ± .20
Gross	.77 ± .09*	Was complied to	.23 ± .21
Paper cutting	.55 ± .16	Refused others	-.03 ± .22
School attendance	-.06 ± .22	Was refused	.61 ± .14*
Failure test		Received friendly	.34 ± .20
Constructive	.78 ± .09*	Gave friendly	.51 ± .17
Emotional	.10 ± .23	Received hostile	.39 ± .19
Intelligence test	.94 ± .03*	Gave hostile	.39 ± .19
Nervous habits	.10 ± .22	*Popularity rankings*	
		Child estimate	.50 ± .17
		Teacher estimate of child popularity	.41 ± .19

*Indicates correlations that are greater than four times their probable error.

consistency from first to final measure was physical attractiveness, which correlated .75 ± .10. The retest on gross motor ability agreed with the former measure by a correlation of .77 ± .09, while paper cutting ability only correlated .55 ± .16 between first and last testing. The constructive score on the failure test correlated .78 ± .09, while the emotional score correlated only .10 ± .23 between the two measurements. The amount of social interactions that a child made seemed to

remain relatively constant, the correlation being .72 ± .11, and the amount of directing the child gave and received showed relationship of .78 ± .09 and .65 ± .13 between test and retest for the two measures respectively. The two measures of intelligence correlated .94 ± .03. The children who were refused by others when the first set of data were gathered seemed to remain high in being refused behavior 9 months later, for the correlation was .61 ± .14.

The measures that seemed to show individual variation from initial to final testing were: 1) initial approach to the constructive-destructive test; 2) attractiveness of personality; 3) routine co-operation; 4) the number of days the child attended preschool the semester the data were gathered; 5) emotional reaction to failure; 6) the amount of communication between child and teacher; 7) being compliant and being complied to in social situations; 8) the amount of refusing others; 9) the number of friendly interactions received and hostile interactions both given and received; and 10) the number of minutes in which nervous habits were manifested. Intensity of social participation and the number of friendly interactions given showed some stability, the correlations being .50 ± .17 and .51 ± .17 respectively between initial and final measures.

The relationships between the initial and final measures of popularity were both less than four times their probable errors; the teacher estimates of popularity correlated .41 ± .19, and the measurements of child popularity correlated .50 ± .17. It must be remembered, however, that the 10 children were not grouped together the second semester as they had been when the initial measures were obtained. The reasons for this were given on page 306.

When correlations were obtained between change in popularity and change in the other measures from initial to final measures, only a few items were found to yield relationships of significance. The following tabulation gives the correlations that seemed to be suggestive of relationship between change in the measure and change in popularity.

Measure	Consistent Popularity	Teacher Estimate of Child Popularity
Intelligence quotient	.42 ± .18	.47 ± .18
Intensity of social participation	.47 ± .18	
Refuses others	.57 ± .16	
Is hostile to others		.45 ± .18

Thus the increase in intelligence was the only measure that tended to show some agreement with change in both measures of popularity. Increase in intensity of social participation and increase in the amount of refusing others showed a possible relationship with increase in popularity as rated by the children. Increase in hostility to others showed some relationship to increase in teacher estimate of child popularity. As the other measures showed only negligible relationships, they have not been mentioned.

Summary of the Relationship Between the Two Sets of Measures on Group D

The measures that showed correlations at least four times their probable error between initial and final testings were: 1) final score on the constructive-destructive test; 2) teachers' rating of physical attractiveness; 3) gross motor ability; 4) constructive reaction to failure; 5) intelligence quotient; 6) the amount other children refused to comply to the child; 7) the amount of communication with children; and 8) the amount of giving and receiving of directions. Measures that showed a possibility of some relationship, but whose correlations were just smaller than four times their probable errors were: 1) intensity of social participation; 2) the amount of friendly behavior manifested; 3) child ranking of popularity; and 4) paper cutting ability.

The measures that seemed to show no relationship between initial and final measurement were: 1) initial approach to the constructive-destructive test; 2) emotional reaction to failure; 3) personality attractiveness; 4) co-operation in routines; 5) school attendance; 6) the amount of communication between child and teacher; 7) the amount the child complied and was complied to; 8) the amount of refusing to comply; 9) the number of friendly interactions; 10) the number of hostile interactions given and received; and 11) the number of minutes in which an oral nervous habit was manifested.

When the change in child estimate of popularity was correlated with the change in each of the other measures, none of the correlations were equal to four times their probable error, but the measures that were greater than three times their probable error were: 1) the amount the child refused others, i.e., the child who increased in popularity tended to refuse others more often; 2) intelligence quotient, the child increasing in popularity also tended to increase in intelligence; and 3) the intensity of social participation, i.e., the child who increased in popularity became more of an active group member.

When the change in teacher estimate of popularity was correlated with increase in the other measures, the two items that were three times their probable errors were: 1) hostility to others, the child who increased in popularity became more hostile; 2) increase in intelligence. Increase in intelligence was the only item that seemed to be related to increase in popularity in both measures of popularity, teacher and child estimate.

SUMMARY AND DISCUSSION OF RESULTS

Summary

The aim of the study was to determine the concomitants of popularity in preschool children.

The subjects were 45 children attending the Iowa Child Welfare Research Station preschools during 1939 and 1940. Group A contained 15 children, Group B, 9 children, and Group C, 21 children. The mean chronological age for the three groups was 44.4 months and the mean intelligence quotient was 131.9. Ten of Group A's subjects were retested at a nine months later date and comprised Group D.

Methods

1. The paired comparison method was employed in obtaining two popularity scores for each child, one from the child's associates and one from two preschool teachers' judgments of child popularity.

2. Data pertaining to the social behavior were obtained from eight ten-minute observations of each child taken during free play periods. The behavioral items observed were: intensity of social participation; and directing, compliant, resistant, communicative, friendly and hostile behavior.

3. Scores for attractiveness of personality, co-operation in routines, and physical attractiveness were obtained for each subject from the combined ratings of two teachers.

4. The following individual tests were given each subject; Ackerman's test of constructive-destructive tendencies; the Revised Stanford-Binet test of intelligence, form "L"; Keister's test of behavior in the face of failure; the ascending and descending of stairs and ladders from McCaskill motor test; and the paper cutting items from Jakway's manual tests.

5. The presence or absence of an oral nervous habit was recorded for each of thirty one-minute observations on each child.

Results Pertaining to Methodology of Popularity Preferences

Findings from the different methods of estimating popularity were: (a) Child and teacher estimates of child popularity seemed to be based upon different criteria. (b) Children tended to favor the last name of the pair in naming their preference. (c) The children showed a slight preference for their own sex, i.e., in bisex pairs the boys favored boys, and the girls favored girls. (d) Teachers estimated boys to be more popular than girls, while the children estimated the sexes to be about equal in popularity. (e) Agreement, although not perfect, was shown between the two teachers' estimates of child popularity in the preschool groups. (f) The total and consistent methods of scoring the child preferences showed high agreement. (g) The two halves of the total method of scoring the preferences showed practically no agreement, although each showed some relationship to the total method. (h) Twenty pairings were repeated after a six-weeks' interval for each of 13 of the subjects of Group C, and when the initial and final rankings of the 5 children upon whom judgments were made were compared only one child was displaced; and (i) When two popularity rankings obtained on Group D, at a nine-months' interval were compared, a positive relationship was found to exist in spite of the fact that the ten children were not in the same preschool subgroups.

Results of the Relationship of Popularity to other Measures for Groups A, B and C

1. Measures that were related to teachers' estimate of child popularity, significant at the 5 per cent level or better, for the combined groups were: (a) the amount and intensity of social participation; (b) the amount of compliant behavior the child manifested and experienced from others; (c) the amount of directing behavior the child manifested and experienced from others; (d) the number of times the child was

refused by others; and (e) the amount of child-child communication.

2. Two measures that were inconsistently related (significant for one group, but not for the combined groups) to both measures of popularity were: (a) the degree to which the child co-operated in preschool routines, and (b) gross motor ability.

3. The following additional measures were inconsistently related to teachers' estimate of child popularity: (a) attractiveness of personality; (b) physical attractiveness; and (c) the amount of hostile and friendly social behavior the child manifested and experienced from others.

4. The following additional measures were inconsistently related to child popularity: (a) constructive behavior in a play situation and in the face of failure; and (b) the amount of child-teacher communication.

5. The measures that were unrelated to either measure of popularity were: (a) paper cutting ability; (b) emotional reaction in the face of failure; (c) the number of days the child attended preschool during the semester of the experiment; (d) chronological age; (e) mental age and intelligence quotient; and (f) oral nervous habits.

6. Unrelated to teachers' estimate of child popularity were: (a) constructive behavior in a play situation and in the face of failure; and (b) the amount of communication between the child and teachers.

7. Unrelated to teachers' estimate of child popularity were: (a) amount and intensity of social participation; (b) attractiveness of personality; (c) physical attractiveness; (d) the amount of child communication; (e) the amount of directing behavior the child manifested and experienced; (f) the amount the child refused others; and (g) the amount of friendly and hostile behavior the child manifested toward others and experienced from them.

Results Pertaining to Group D

1. The following measures showed relationship (the rank difference correlations were four times their probable error) between initial and final measures: (a) final score on the constructive-destructiveness test; (b) physical attractiveness, as rated by teachers; (c) gross motor ability; (d) constructive reaction to failure; (e) intelligence quotient; (f) the amount other children refused to comply to the child; (g) the amount of communication with children; and (h) the amount of giving and receiving of directions.

2. The items that showed possible relationship (the correlations were .50 $\pm$ .17 or better, but less than four times the probable error) were: (a) intensity of social participation; (b) the amount of friendly behavior the child manifested towards others: (c) child ranking of popularity; and (d) fine motor ability.

3. When change in popularity was related to change in the other measures over a nine-months' interval, none of the correlations were four times their probable error. However, the following measures seemed to show some relationship (the correlations were just smaller than four times their probable error): (a) increase in intelligence seemed related to increase in both measures of popularity; (b) change in teachers' estimate of child popularity seemed positively related to manifestation of hostile behavior, i.e., the child increasing in this measure of

popularity also increased in the number of hostile interactions he manifested toward others; and (c) the child increasing in child popularity also increased in the amount he refused others and increased in the amount of social participation.

Discussion of Results

When judging popularity in children, adults are likely to base their estimates upon behavioral criteria. The teachers of groups A, B and C seemed to judge the socially active child as popular and the isolated child as unpopular. Certainly to be popular a child needs to be a participant of a group, but the active member of a group need not necessarily be the popular member. The dictatorial leader might hold the group forcefully or by sufferance, while another child might be the real center of the group because of his attraction for others. The latter type of child is likely to be more popular than the former. Thus it is to be expected that children's statement of preferences and teachers' estimate of child popularity would not show high agreement.

If any one factor can be said to have been indicative of child popularity, as far as the groups of the present study and that of Koch's are concerned, the item, co-operation in routines, seems to have been such a factor. Analysis of this scale reveals that it contained many items pertaining to adjustment and the degree to which the child entered into the situation. Items pertaining to quick, unobtrusive adjustment had a heavy weighting, while items of refusing, dawdling, silly acts, and so forth, were given low scores. Thus it seems that quick adaptation to the situation without making a disturbance is typical of the most popular child, and the reverse of this is true of the unpopular child.

The results obtained on Groups A, B, and C and that of Koch's (9) seem to raise some suggestions for further study.

1. Might not a measure of ability to adapt in social situations be one of the best measures for future investigation of popularity?
2. To interpret child popularity one needs to know more of the criteria upon which such preferences are based. Might not direct questioning of the child as to why a popular child was preferred lead to some elucidation of child values of the popular child?
3. Might not further research along the line of relationship of change in popularity to change in other measures yield information helpful to the guiding of the unpopular child?

REFERENCES

(1) Ackerman, N. W.: Constructive and destructive tendencies in children. Am. J. Orthopsychiat., 1937, 301-319

(2) Ackerman, N. W.: Constructive and destructive tendencies in children: An experimental study. Am. J. Orthopsychiat., 1938, 8, 265-285.

(3) Carr, Virginia: The social and emotional changes in a group of highly intelligent children during a program of increased educational stimulation. Master of Arts thesis, University of Iowa, 1938.

(4) Eesley, G.: A study of teacher's ratings of the attractiveness of personality and physical attractiveness of nursery school children. Merrill-Palmer School Unpublished Study, 1933. Accepted as a thesis for the degree of Master of Arts, Wayne, University.

(5) Goodman, M.: The effect of training children to face difficult situations. Master of Arts thesis, University of Iowa, 1938.

(6) Hsia, J. C.: A study of the sociability of elementary school children. Teach. Coll., Columbia Univ., Cont. to Educ., 1928, No. 322. Pp. 1-64.

(7) Jakway, I. A.: Manual of achievements of preschool children. Unpublished Master of Arts thesis, University of Iowa, 1938.

(8) Keister, M. E.: The behavior of young children in failure. Univ. of Iowa Stud., Stud. in Child Welfare, Stud. in Preschool Educ., 1937, 14, 28-82.

(9) Koch, H.: Popularity in preschool children. Child Develop., 1933, 4, 164-175.

(10) Lippitt, R. N.: Popularity among preschool children. Unpublished Ph.D. thesis, University of Iowa, 1940.

(11) McCaskill, C. L.: A study of common motor achievements at the preschool age. Master of Arts thesis, 1936.

(12) Olson, W. C.: The measure of nervous habits in normal children. Minneapolis, Minn.: University of Minnesota Press, 1929. Pp. xii, 97.

(13) Roberts, K. E., and Ball, Stutsman: A study of personality among young children by means of a series of rating scales. J. Genet. Psychol., 1938, 52, 70, 79-149.

(14) Thurstone, L. L.: An experimental study of nationality preferences. J. Genet. Psychol., 1928, 1, 405-424.

(15) Updegraff, R., et al: Practice in Preschool Education. New York, McGraw-Hill Co., Inc., 1938. Pp. xvi, 408.

THE LOCATION OF THE SUBORDINATE CLAUSE IN ORAL AND WRITTEN LANGUAGE[1]

EDITH A. DAVIS

Now and then our careful quantitative methods fail to bring to light a fact which is not only important, but almost self-evident - once it has been pointed out. American students of language development are indebted to Fritz and Grace Heider (4) for calling to our attention a characteristic of certain types of subordinate clauses, first noted by William Stern. The Heiders remind us that there is a special difficulty involved in placing a subordinate clause at the start of a sentence, interrupting the main thought by a qualifying statement, or expressing a possibility, and they suggest that here may lie a partial explanation of the differences in usage between hearing children and older, educationally retarded deaf children, when asked to write a description of a motion picture.

Since in my own analyses (3) I had not studied the position of the subordinate clause in the sentence, I recently tabulated from this point of view the subordinate clauses occurring in 18,650 remarks of children recorded in a standardized situation. Of the subjects, 248 were children enrolled in kindergarten, who were within a few weeks of 5½ years old, and 125 were within two months of 9½ years old. Fifty remarks were recorded for each child, while he was engaged in play with an assortment of miniature Wild West toys in a room alone with me, in the school building he attended in Minneapolis or St. Paul, Minnesota.

These children were younger than Heider's, and oral language of course tends toward loosely constructed or fragmentary sentences. Sometimes the main clause was implied in a remark or question of the examiner, and there was no way of knowing whether in the child's mind it came before or after the subordinate clause. The findings are therefore not entirely comparable with those of the Heiders. It seems clear, however, that clauses of time and condition usually introduce or interrupt the main clause. This is true of a small percentage of relative clauses, and occurs occasionally in an object or other type of clause. That the construction is associated with a high level of linguistic development is shown by the fact that it is used more often by the older children, and that it is the brighter kindergarten children who use it at all.

Thus, of 65 children whose Pintner-Cunningham IQ's were 90 or below (26 per cent of the 248 kindergarten children), 14 or 21 per cent placed one or more clauses early in the sentence, while 16 or 25 per cent used no subordinate clauses at all. Of 41 children (16 per cent of the group) whose Pintner-Cunningham IQ's were 120 or above, 19 or 46 per cent placed clauses where they preceded or interrupted the main clause, while only one, a shy little girl with a serious speech difficulty, used no subordinate clauses. The mean IQ according to this test was 92 for those using no clauses, 103 for those placing clauses at the end of the sentence, and 108 for those placing the subordinate clause where it preceded or interrupted the main clause. The mean usage for the brightest children was 1.74 times per child, and for the entire group, 1.53 times.

[1] *From Idaho Child Welfare Services, Boise, Idaho.*

Further proof was found by inspecting individual records of a few children whose outstanding ability has been established by repeated psychometric tests given at the Institute of Child Welfare of the University of Minnesota and by their academic progress since the language records were obtained in 1930-1933. A boy who has consistently tested at about 170 placed a subordinate clause at the start of four of his sentences, while a girl of similar high ability used the construction three times.

This relationship with intelligence was less clear-cut in the older group of children. Scores on the Pressey Intermediate Classification Test were available for 112 of the 125 subjects. Of these, 15 or 13 per cent scored at 100 or below and 15 or 13 per cent at 130 or above. These limits were used because this test gave a mean IQ for the entire group of 115.6. Twenty-seven per cent of the lowest 15 but 54 per cent of the highest 15 placed clauses early in the sentence. Eight children used no subordinate clauses at all. The mean IQ for these was 96.6, while that for the 62 who used the construction was 118. The mean IQ for the remaining 42, who used clauses, but placed them at the end of the sentence, was 115. That there is a relationship with age is shown in Table 1. Although even at 5½ years there is a definite tendency to

TABLE 1

PER CENT OF CHILDREN PLACING THE SUBORDINATE CLAUSE WHERE IT PRECEDES OR INTERRUPTS THE MAIN CLAUSE

Age in years	Number in group	Number placing clause early	Percent	Mean per child in group	Mean per child using	Mean use by brightest	Use per 1000 words
5½	248	73	29	0.45	1.53	1.74	1.97
9½	125	67	54	1.33	2.48	2.75	4.10

place clauses of time and condition toward the beginning of the sentence, nevertheless this is done in a larger percentage of instances by the older children. This is shown in Table 2.

We may conclude, then, that subordinate clauses of time and condition are usually found early in the sentence; that the 5-year-old child has already learned this usage, especially if he is advanced intellectually; and that the pattern is more firmly fixed in the child of 9 than in the child of 5. The total percentage of time clauses coming early in the sentence, however, even for the older children, is considerably smaller than the 90 per cent reported by Heider. We cannot say with assurance whether this difference results from further changes with age, or is a characteristic difference between oral and written language, although Bushnell's (2) study throws a little light on this point.

Bushnell (2) published 15 pairs of oral and written compositions by tenth grade pupils on the subjects "How I learned a lesson" and "A good plan that went wrong". These were selected as representative pairs from a total of 100 pairs that constituted the material of his study. Pupils did not know that their oral themes were being taken down by a stenographer. Twelve days afterward the same topic was assigned for a written theme, with a time limit of 20 minutes. I tabulated the subordinate clauses in these 15 pairs of themes, and secured results as shown in Table 3.

TABLE 2

POSITION OF VARIOUS TYPES OF SUBORDINATE CLAUSES
WITH REFERENCE TO MAIN CLAUSE

	Time		Condition		Relative		Others
	Precedes	Follows	Precedes	Follows	Precedes	Follows	Precedes
			(Age 5½ years)				
Number	38	47	22	9	16	261	36
Per cent	45	55	71	29	6	94	
No. per child	.15	.19	.09	.04	.06	1.05	..14
No. per 1000 words	.67	.83	.39	.16	.28	.61	.63
			(Age 9½ years)				
Number	64	49	54	17	22	154	26
Per cent	57	43	76	24	12.5	87.5	
No. per child	.51	.39	.43	.14	.18	1.23	.21
No. Per 1000 words	1.58	1.21	1.33	.42	.54	3.80	.64

TABLE 3

POSITION WITH REFERENCE TO MAIN CLAUSE OF VARIOUS TYPES OF SUBORDINATE
CLAUSES IN ORAL AND WRITTEN COMPOSITIONS OF BUSHNELL'S 15 SUBJECTS

Type of composition		Type of clause							
		Time		Condition		Relative		Others	
		Precedes	Follows	Precedes	Follows	Precedes	Follows	Precedes	Follows
Oral	Number	21	6	4	1	2	38	1	134
	Per cent	78	22	80	20	5	95	1	99
	Mean	1.4	.4	.27	.07	.11	2.53	.07	8.20
Written	Number	25	17	6	3	6	39	4	104
	Per cent	60	40	67	33	13	87	4	96
	Mean	1.7	1.1	.4	.2	.4	2.6	.27	6.93

In so far as conclusions are warranted on the basis of so small a sampling, the position of clauses within the sentence is the same for oral and for written language. This was Bushnell's main conclusion from the study as a whole - the merits of oral expression "are of practically the same sort but not so great as those of written language," since the oral products were "more fluid, unorganized, and incoherent" than the written. On the whole, they were somewhat longer. Since they were recorded in a formal class-room situation they are far from comparable to the casual, give-and-take conversation which made up my own data. Nevertheless, Bushnell's material is a unique means of comparing spoken and written language. We have too often judged the speech of children by standards based on the published writings of adults.

The index of subordination (5) was slightly higher in the written than in the oral themes. This finding also is of value, although based on only 15 pairs of themes, because it constitutes a definite link in the developmental sequence which is being established (1,3,4,5), as is shown in Table 4. The index rises with age and educational experience,

and is higher in written than in oral language, although it is affected by the situation and by the type of thought process involved. For Bushnell's 15 subjects, there was little relation between the subordination index and the excellence of the compositions as rated by the Van Wagenen Scale. The correlation coefficient by the Spearman rank-differences method was .12 for written, -.39 for oral compositions.

TABLE 4

THE INDEX OF SUBORDINATION AT SUCCESSIVE DEVELOPMENTAL LEVELS, IN ORAL AND WRITTEN LANGUAGE, AS REPORTED BY VARIOUS INVESTIGATORS

Oral				Written		
Author	Type of group	Number of cases	Index	Author	Type Cases	Index
Davis	Age 5½ years	248	.10	LaBrant	A Grades 4-8	482 .28
	" 6½ "	63	.12		B " 9-12	504 .36
	" 9½ "	125	.17		C Psychologists	21 .46
				Anderson	Univ. students	111 .50
Bushnell	Tenth grade	15	.34	Bushnell	10th grade	15 .39
				Heider	Hearing age 8-14	817 .10-.17
					Deaf age 11-17	301 .07-.12

Table 5 shows that the use of infinitives and participles, or verbids (6), as well as of subordinate clauses, is slightly greater in the written compositions, although the mean number of independent predicates per subject is greater in the oral compositions. All these minor findings substantiate Bushnell's contention that there is better organization of written than of oral language.

TABLE 5

USE OF FINITE VERBS AND VERBIDS IN ORAL AND WRITTEN COMPOSITIONS BY BUSHNELL'S 15 SUBJECTS

Mean number per subject	Oral	Written
Independent predicates	27.7	20.9
Dependent predicates	13.1	13.6
Infinitives	4.7	5.3
Participles and gerunds	1.3	2.1

Although Heider's hearing subjects used complex types of sentence structure to a greater extent than the deaf subjects, the deaf children did make extensive use of two types of subordinate clause, the object clause introduced by "that" and the causal clause. Analysis shows that these two clauses are almost always placed at the end of the sentence, and therefore do not present the special difficulty met with when the thought of the main clause is interrupted. The deaf children were much more likely than the hearing to express the connective "that" in their object clauses and to use "which" in relative clauses where hearing children would use "what", "that", or omit the relative pronoun altogether. Heider points out that a choice of connectives, the use of a clause to modify a particular word rather than the sentence as a whole, and the introduction of an element of doubt or possibility as with the interrogative in indirect discourse, and especially in the object clause introduced by "if" or "whether", greatly enhances the difficulty of expression. The deaf were found to follow a rigid, fixed pattern of construction while the hearing children used a more flexible mode of expression.

Of my own subjects, children at both age levels tended to express the connective in relative clauses and to omit it in object clauses, although the older children used the elliptical form less than the younger ones. This is shown in Table 6. Of the relatives expressed, "that" is used much more often than "who", "which", or "what".

TABLE 6

AGE DIFFERENCES IN EXPRESSING CONNECTIVES

	Object clauses with "that"		Relative clauses	
	5½	9½	5½	9½
Number expressed	8	26	134	120
Per cent	3.5	11	48	68
Number omitted	218	206	143	56
Per cent	96.5	89	52	32
Use per child	.91	1.86	1.12	1.41
Use per 1000 words	3.99	5.73	4.90	4.34

Since the usage of relative clauses is less in proportion to total words used for older than for younger children, it appears that hearing children of 5½ years do not experience serious difficulty in using clauses to modify a single word rather than a whole clause. There is an increase both per child and per words used in the incidence of object clauses with "that".

Another point discussed by Heider is the use of the interrogative in indirect discourse. This also has been neglected in our studies of language. The difficulties in managing an object clause of this kind are apparent. Compare the following sentence, an actual recording from a girl of 5½ years, IQ 110, with the corresponding form in indirect discourse:

I said to the teacher, why don't you send to the principal, there's so much noise here.

I asked the teacher why she didn't send to the principal, because there was so much noise.

Again, quoting from the chatter of an only girl of the same age, IQ 101;

I asked Mother if I could go up to Grandma's with her in the morning, she had to plan her dress, and she said, No, you have to go to school, but you can go in the afternoon.

In spite of its difficulties, my younger subjects used indirect discourse rather more than did the older children. There were 208 instances, 3.67 per 1000 words, for the former, and 87, or 2.15 instances per 1000 words, for the latter. Heider's analysis found the interrogative in indirect discourse used with equal frequency by deaf and hearing children. Bushnell's 15 subjects used it 26 times in their oral themes and 11 times in the written ones, which was 13.3 and 5.5 per cent respectively of all the subordinate clauses.

It seems obvious that as a result of Heider's study, workers with children's language will scrutinize more carefully than ever the mechanics of sentence structure. Certainly we have by no means exhausted the possibilities for new approach to research in this most fascinating field.

REFERENCES

(1) Anderson, J. E.: An evaluation of various indices of linguistic development. Child Develop., 1937, 8, 62-68.

(2) Bushnell, P. P.: An analytical contrast of oral with written English. Teachers' Coll. Contrib. To Educ. 451, Bur. Publ., Columbia Univ., New York, 1930.

(3) Davis, E. A.: The development of linguistic skill in twins, singletons with siblings, and only children from age five to ten years. Univ. Minnesota Inst. Child Welf. Monog. No. 14, 1937.

(4) Heider, F., and G. Heider: A comparison of sentence structure of deaf and hearing children. Chapter IV in Studies in the Psychology of the Deaf, Psych. Monog. 52, No. 1, 1940. Am. Psych. Assoc., Columbus, Ohio.

(5) LaBrant, L. L.: A study of certain language developments of children in grades 4-12 inclusive. Genet. Psych. Monog., 1933, 14, No. 5.

(6) Rankin, J. R.: A new plan of English grammar. Henry Holt and Co., New York, 1933.

ANTHROPOMETRIC AND ORTHOPEDIC EXAMINATIONS[1]
A Technique for Use with Children

MABEL POTTINGER BAUM
and
VERNETTE S. VICKERS

In 1930 the Department of Child Hygiene of the Harvard School of Public Health instituted the taking of certain physical measurements on a group of children as a part of a study of physical growth and development (4). Certain of these measurements are of value to the pediatrician and may be taken by him; others are of interest in the more intensive study of children, such as is conducted in child study centers. In either case, the standardization and adaptation of techniques is of great importance.

On each visit to the Center the children have between seven and eleven separate examinations, of which the orthopedic and anthropometric examinations are but two. The time allowed for these examinations and the freedom of action of the child are both necessarily limited. It is the problem of the examiner and assistant in this department to collect the wanted information in thirty-five to forty minutes with the maximum efficiency and minimum of fatigue and boredom for the child. It is the hope that some of our solutions of this problem and some of the conclusions may be of interest to others engaged in similar work.

We have come to feel that the technique of taking body measurements and making observations of posture covers a great deal more than skill with calipers or a knowledge of body mechanics. It implies a consistent and well thought out plan of procedure, a cooperative and trained personnel, and an appreciation of the fact that the examination must be made to offer something to the child besides a period of restriction.

For most efficient operation, the examination must be considered as a whole and each step planned for the utmost conservation of time and motion. When the plan has been worked out, it should be adhered to until the mechanics of its order become almost automatic. Only then is the examiner's mind unencumbered by the routine, leaving him free to talk to the child, keep him interested and obtain his cooperation. It also means better observation in a thoroughly easy atmosphere and makes possible a greater sharing of responsibility between examiner and assistant. It is believed that the assistant should be a great deal more than a mere recorder. The more nearly his training, interest and responsibility approach that of the examiner, the better the examination, the cooperation and the records.

Training in recognizing and handling the psychological problems of small children is of as much importance as experience in the special skills of the examination. The technique in handling the child must aim at assuring reasonably accurate measurements and typical posture for an individual child. To a very large extent, whether or not a child is quiet and cooperative depends upon the adjustment of the routine to avoid fatigue, boredom or fear (insecurity, fear of apparatus, fear of falling

[1] *From the Department of Child Hygiene, Harvard School of Public Health.*

from high tables, etc.). The child should leave with a pleasant memory and a willingness to return for his next routine examination, for the life of a longitudinal study is dependent upon cooperative mothers and children.

The children in the group range in age from 6 weeks to 11 years. From the point of view of handling, they fall into four large groups, with special problems in each case. The infants who are brought into the examining room quiet and well wrapped resent being stripped suddenly of all protective coverings. These infants are left on their blanket and only partially exposed at first. Certain measurements, especially of the head, which restrict the child's movements, tend to antagonize him. These measurements are postponed until near the end of the examination. The usual method of measuring total length is to hold the head and trunk, with the head against the fixed upright, while the movable arm is brought against the soles of the feet. If this process is reversed, much less opposition is offered. The assistant holds the feet, while the measurer brings the movable arm to the vertex with one hand and needs only to give momentary and light restriction at the chin.

The second group contains the youngsters from 15 to perhaps 24 months. With these children, language, as such, is difficult. Above all, they hate to leave their mothers and do not like to be lifted high off the floor and placed on a table. Their reaction is a combination of fear and resentment. This is the age period during which it is most difficult to secure cooperation. A child of this age cannot be hurried if he is to accept the examination of his own free will. He must be made to feel friendliness around him, a willingness on the part of the examiner to let him play with the toys and to show him how they work. A little of this and he will sit on the examiner's lap and allow his feet to be examined. Finally, he will sit on the table, usually under some protest but without fear and without the tantrum with which he would have met being lifted onto the table within his first ten seconds in the room, as may be done with slightly older children.

Into the third group fall the majority of the children at the present time. For the most part, all they want are toys and talk, but some of them are afraid, miss their mothers, or are tired or unhappy. Sometimes these emotional upsets can be resolved completely, sometimes only partially, but fear and insecurity must be met if the child is to be measured with any degree of accuracy or is to be observed in anything like his normal condition (3).

The fourth group is made up of the older children who have reached the third or fourth grade. For the most part, these children have become so accustomed to the routine at the Center that they accept the removal of their clothes as a matter of course. However, there are a few whose enjoyment or dislike of the whole procedure is entirely conditioned by their embarrassment at being without clothes. We have found that by allowing the child a very scant pair of pants, a tense, unhappy child can be changed into a pleasant visitor. The pants do not interfere either with measurements or with observations of body contours; the lack of them does.

Considerable thought has been given to the selection of toys which can be held in the hands of very small children. They cannot be toys

that the child needs to put down on the floor to manipulate to the best advantage. They must have a value for the child alone and not need a second person to make them entertaining. It is necessary that they not be noisy and that they have a wide age range of interest and be durable, washable and colorful. At present our collection consists of a string of "plakies" (colored plastic discs on a chain), a transparent celluloid ball containing water and a toy fish, a doll to be slipped over the hand and manipulated, and a small "pigs-in-clover" puzzle. These toys, though simple, seem to have the qualities which make them adaptable for use while the child is on the table. They allow the child to entertain himself, offer an introduction to conversation, and do not interfere with observation and measurements. Their interest range seems to have no age limits, and the children do not tire of them quickly.

A different type of toy is available after the child has been examined on the table and is required to stand for some length of time for posture examination and measurements of body angles. We have an aquarium of colorful tropical fish and snails at approximately the child's eye level, a Russian doll with a graduated series of smaller dolls inside, and an apple of the same arrangement. There is adequate space in front of the fish tank for the use of these toys, so that the field of the child's play operations is limited to the satisfaction of both the examiner and the child. There is no necessity for putting a conscious strain on the child by having to tell him where or how to stand.

The practice has been made of not allowing the children to play with the measuring instruments, which are expensive and must be exact. We are perfectly willing that the child examine them, if there seems to be any concern about their use, but it is believed that turning over our tools to the child adds nothing to the child's feeling of security and frequently leads to a kind of disrespect for the whole of the examination procedure. We have tried both methods and find ourselves in disagreement with Dr. Boyd who says, "In taking such a long series of measurements on young children, with repetitions at monthly intervals, it is necessary to obtain their confidence... We permit the children to investigate all the equipment, a cooperative child being considered more important than expensive calipers, and we use no extraneous toys to entertain them"(1).

In carrying out the examination, an effort has been made to get away from the method of starting at the child's head and working down to his feet. This, we feel, has increased the ease and efficiency of the examination and enables us at all times to preserve the optimum conditions for each measurement and observation. Taking the child's length after he has been on the table for ten or fifteen minutes results in a nearer approach to his maximum length. Taking his standing height immediately after he leaves the table allows a minimum loss of height, and followed immediately with the posture examination, a minimum of postural fatigue.

The child usually removes his bathrobe spontaneously when he comes into the examining room and is lifted or climbs onto the table, where he sits with his legs hanging over the edge. The examiner chats with him while looking at his extremities, measuring his head, etc. He is allowed to sit in this position, asking and answering questions. After he has become used to the height of the table, the examiners and the apparatus,

he lies down, usually feeling secure and unafraid. He is consequently relaxed, which is an important contribution to the accuracy of the measurements. The measurements that are taken and the observations fall into the following groupings:

Sits on the edge of the table with the legs over the side:
1. Head breadth and depth measure. Face measured.
2. Arm and leg muscles estimated and handedness determined.

Swings feet around and sits with legs extended:
1. Sitting observed.

Lies on back:
1. Chest and iliac measurements taken.
2. Observations of chest.
3. Body angles measured.

Lies on side:
1. Chest depth measured.
2. Lateral structures tested.

Lies on stomach:
1. Hip extension measured.
2. Anterior structures tested.

Rolls over onto back again:
1. Length and several other measurements taken.

Child is lifted to the floor:
1. Standing height and other measurements taken immediately. (Care is taken not to let the adventurous children jump off the table for fear of spoiling the measure of maximum height.)

Walks the length of the room:
1. Gait observed.
2. Feet measured.

Stands in front of the fish tank - plays with toys, if he wishes:
1. Posture observed.
2. Body angles measured.
3. Feet observed for pronation, knock knees, etc.
4. Lateral contours of back observed.

"Three Stunts:"
1. Fingers to floor.
2. Tipped over backwards to test abdominal muscles (head in assistant's lap).
3. Tipped forward, hands on the floor, to test back muscles.

The whole examination takes from 35 to 40 minutes. If an interruption occurs, the child is asked to sit down. The period of standing is interrupted by giving the child a chance to walk about the room. This schedule is seldom varied, and we have found that it results in a minimum of fatigue and a maximum of cooperation.

All measurements are carried out by each examiner and are recorded separately. It has always been required that the measurements of the

two examiners check within 5 millimeters, and during the past 4 years measurements of the head have been required to check within 3 millimeters. Measuring infants and young children accurately is a difficult procedure at best, and the technique of holding the child in proper position is most important. The individuals who take the measurements must be well trained and adept in handling children, and it is desirable that the same persons work together constantly, so as to develop cooperative techniques.

In the choice of measurements, those included in the International Agreement (4) have been adopted, in so far as possible, and unless otherwise noted the techniques prescribed in this agreement have been followed. Where recumbent, rather than erect, measurements have been taken, the same landmarks have been used, and an attempt has been made to have the other features of technique as comparable as possible. Unless otherwise stated, the measurements described have been taken in the dorsal recumbent position and have been applied at all ages.

MEASUREMENTS INCLUDED IN THE ANTHROPOMETRIC STUDIES

1. WEIGHT. This is measured in pounds and ounces on the infants' scale, and in pounds, half-pounds and quarter-pounds on the children's scale. Weight has been measured independently of the other measurements by a different person.

2. TOTAL LENGTH. This measurement is obtained by placing the child in the dorsal recumbent position, with the soles of the feet against the broad, right angle plate forming the base of the anthropometer. The length is read at the horizontal bar of the anthropometer when the latter is placed against the vertex.

3. SITTING LENGTH. This measurement is made with the head and trunk in the dorsal recumbent position and with the buttocks placed firmly against the fixed, right angle plate. The hips and knees are flexed to right angles. The length is read to the vertex, as in the total length measurement.

4. STATURE OR STANDING HEIGHT. The child stands with his heels, buttocks and occiput against the vertical upright, unless insistence upon contact at all of these points causes marked change from natural posture. The eyes are horizontal and in approximately the same plane with the external auditory canals. The height is taken to the vertex, using the right angle wooden arm of the anthropometer.

5. SITTING HEIGHT. The child sits on the horizontal seat of the vertical measuring device, with his knees flexed to a right angle. The sacrum, dorsal trunk and back of occiput are against the upright anthropometer, and the eyes are horizontal. The measurement is taken to the vertex, as in standing height.

6. HEAD LENGTH. This measurement is the maximum glabello-occipital diameter of the vault. It is taken from the most prominent point of the glabella to that point on the vertical line bisecting the occiput which shows the maximum spread of the calipers.

7. HEAD BREADTH. This measurement is the greatest transverse diameter which can be found in a horizontal plane above the supramastoids and zygomatic crests and is taken with the spreading calipers.

8. HEAD CIRCUMFERENCE. This measurement is taken with a linen tape

in the horizontal arc which passes through the most prominent part of the occiput and just above the supraorbital ridges.

9. HEAD HEIGHT. This is a computed measurement and does not correspond with the "Height of the Head" of the International Agreement. The distance from the soles of the feet to the superior border of the auditory opening, taken on the horizontal anthropometer, is subtracted from the total length.

10. CHEST BREADTH. This measurement is taken in the horizontal plane at the level of the base of the ensiform cartilage. Spreading calipers are used, and the reading is made in mid-respiration.

11. CHEST DEPTH. This measurement is made in the same plane and with the same technique as chest breadth.

12. CIRCUMFERENCE OF CHEST. This measurement is taken with a linen tape in the same plane as chest breadth at mid-respiration.

13. BI-HUMERAL DIAMETER - ERECT. This is the maximum distance between the two prominences formed by the deltoid muscles when the child is standing. Sliding, wooden calipers are used.

14. BI-SPINAL DIAMETER. This is the distance between the two anterior superior iliac spines. The technique used has differed from that adopted in the International Agreement in that spreading calipers are used, and points are applied to the inferior surface of the spines rather than to their summits.

15. BI-CRISTAL DIAMETER - RECUMBENT. This is the maximum distance between the external margins of the iliac crests, using spreading calipers.

16. BI-CRISTAL DIAMETER - ERECT. The maximum distance between the iliac crests is measured with the sliding calipers.

17. BI-TROCHANTERIC DIAMETER - ERECT. The maximum distance between the external surfaces of the great trochanters.

SUMMARY

1. A planned technique of examination is necessary to assure accuracy in measurement and to eliminate repetitive movements and unnecessary changes in position. A team of examiners aids in the efficiency of the procedure.

2. Finesse should be cultivated in handling different age groups.

   A. Infants need quiet, confident handling and protective coverings.

   B. Children between 15 months and 2 years should be offered toys, given verbal assurance and have their movements restricted as little as possible.

   C. Preschool children usually need only toys and conversation. Occasionally, a child presents an exterior calmness which covers an emotional tenseness. The examination should be interrupted for these children and an attempt made to remove the cause of the fear or worry.

   D. School children appreciate conversation on contemporary interests. These include discussions of movies, radio programs, clubs, play activities, current history and, of course, varies with the intelligence, sex and age of

the child. At this age a pair of snug pants worn part of or the entire time adds much to the happiness of the child.

3. Toys should be limited in number, have entertainment without nuisance, value and appeal to all ages. Measuring instruments should not be included among the playthings.

REFERENCES

(1) Boyd, Edith: The experimental error inherent in measuring the growing human body. Am. J. Phys. Anthrop., 1939, Vol. 13, No. 3, Oct.-Dec.

(2) Hrdlička, Aleš. Anthropometry. Philadelphia, Wister Institute of Anatomy and Biology, 1920.

(3) Shirley, M., and Poyntz, L. Children's emotional responses to health examinations. (To be published)

(4) Stuart, H. C., and Staff. Studies from the Center for Research in Child Health and Development. I. The center, the group under observation, and studies in progress. Monog. Soc. Research Child Develop., National Research Council, Washington, D. C. Vol. IV, No. 1,Serial No. 20, 1939.

DEVELOPMENT AND CULTURAL PATTERNING IN CHILDREN'S PROTESTS

MARY SHIRLEY[1] AND LILLIAN POYNTZ[2]

In the care and training of children adults are always imposing routines and procedures that children dislike and against which they object as forcefully as they are able. Simultaneously adults are always imposing a code for social conduct that is approved by the particular culture in which they live. Our culture demands a considerable degree of emotional fortitude and disapproves of persons that give vent to their feelings of fear and pain in outcries and expressions of distress. Parents and other adults and even children inculcate an attitude of stoicism toward hurts and minor ills in the very young by such precepts and taunts as "Be a little man," "Grin and bear it," "Don't be a cry-baby," and "You're a sissy." By the nature of their protests children reveal not only developmental trends in their use of speech for the purpose of controlling adult behavior, but also a growing conformity to the social code of moderation in the expression of emotions.

The protests that are analyzed herein were made during various types of physical examinations to which more than 200 children have been subjected at half-yearly intervals over a period of years, as part of a long term growth study.[3] None of the procedures involved real discomfort or pain to the child, with the possible exception of the finger prick for obtaining a small sample of blood and the occasional innoculations. All the examiners were highly skilled in the techniques they were using and devoted much attention to reassuring the child and obtaining his cooperation. At all ages there were a number of children who made little or no objection to the examinations. Also it was impossible for examiners who were busily engaged in making and recording a series of observations to record verbatim all the child's remarks. Hence this analysis is based on a qualitative study of the different types of protests made at different ages rather than on the frequency of their use.

In the accompanying table the different protests are listed separately for each sex at the age of their first use. A developmental trend in these protests is unmistakable. The earliest verbalized protest is the call for Mamma or Daddy on whom the child is accustomed to call for aid. Since neither parent is present at the examination this type of protest is non-adaptive and non-specific for the situation. Both boys and girls use this type of protest from the time they begin talking to the age of 5 years, but the frequency declines with age. Almost equally unadaptive in this situation is the protest I want to go home. Such protests are perhaps a proof from the lips of the child himself that parents and home are the basis of whatever security he feels.

[1]*School of Social Work, Smith College.*

[2]*Supervisor W.P.A. Nursery Schools, Mobile, Alabama.*

[3]Data were collected at the Center for Research in Child Health and Development, Harvard School of Public Health. For a description of the examinations see Harold C. Stuart and Staff: *Studies from the Center for Research in Child Health and Development, School of Public Health, Harvard University. I. The Center, the group under observation, sources of information, and studies in progress. Monog. Soc. Res. Child Develop., National Research Council, Washington, D. C., Vol. IV, No. 1, Serial No. 20, 1939.*

TABLE 1

VERBAL PROTESTS AT EXAMINATIONS

Age	Boys' Protests	Girls' Protests
2 years	"Mamie! "Sistah!" "All done?" "Go out."	"Mamie!" "See Daddy?" "No." "All done?"
2½ years	"I wanna see Mamma!" "I want to go home!" "Daddy come soon?" "Shoes back on?" "I don't wanna! " "Don't do it again!"	"I want my Daddy!" "Wanna go home!" "No more!" "That hurts!" "Down!" (from examining table) "Where is the doctor?" "I wanna go upstairs." (to playroom) "I wanna wet again."
3 years	"No more! No more!" "Don't!" "Get dressed." "You go with me." (begging assistant to accompany him to examination) "I might fall." (off examining table) "I'm not going to get on the table. If I do, I'll cry."	"I don't want to get on the table." "I wanna get down!" "I wanna get my clothes on right now!" "You're not going to take my shoes off!" "Don't stick the needle in me!"
3½ years	"No! No!" "No, I won't." "I don't wanna do it any more!" "I don't want my pictures taken." "I don't want my clothes off." "I want to go downstairs, and I want my clothes on." "You aren't going to take my shoes off, are you?" "Put my clothes on!" "Can I get down? I don't like that." "I don't have to rest, do I?" "Don't talk to me!" "I'm not going to cry!"	"I want my mamma now!" "I've got to go home." "When am I going home?" "I don't want to be measured." "Am I goner (going to) get my finger (pricked)?" "I'm not going to cry!" "It's goner hurt." "Don't do it again."
4 years	"Where's Mamma?" "Don't hurt me!" "No pricks!" "I don't need to (be measured). I don't like it in there. "My Mother said not to take my shoes off!" "I don't like it." "Now don't put one of them in me." "That's gointa hurt."	"No more!" "I don't like this!" "I don't want you to." "Don't let me fall!" "I can't lie on my back." "I don't cry!" "I'm not going to cry. They're just going to measure me." "The tears keep coming, but I keep wiping them away." "I want to sit in your lap." "Stay with me." "Oh, my gosh!" "Ow!"
4½ years	"I want to get out of here right now!" "I'm going home in a little while!" "Don't let it (instrument) touch me!"	"Don't let her do me." "What's that for?" "It didn't hurt me, but I don't like the taste of it." (dental wax) "Will you hold my hand?"
5 years	"You're going to make me have my picutres taken!" "Do I get a vaccination?"	"Am I going to have my finger pricked?" "I didn't cry this time!"
5½ years	"You don't have to cut that in?" "No it's nasty. (dental wax) I don't want to." "The bottom one hurts 'cause he pushes it down. " (dental impression)	"I don't want to get my teeth fixed today." "It tastes awful, but I like it." (dental wax) "Oh, I'm scared to death."
6 years	"I wanna play." "I wanna leave my underwear on." "I was choking. Is that all?" "I was just afraid it was going to hurt but it didn't." "It's a needle just the same." "Anybody that cries is a sissy."	"Any teeth today?" "I don't want to get vaccinated." "I'm goner tell my mother. You can't do that to me!" "I was a little bit scared. I cried last time."
6½ years	"Don't hurt me!" "Don't put it in too long." "Don't open my mouth too wide." "I hope I don't get that stuff." "Are you going to take a tooth out?"	"Don't pull out my loose tooth!" "I thought it was going to hurt, but when I sat in the chair I said to myself 'It's not going to hurt!' and it didn't."
7 years	"Oh!" "Ow!"	

The second type of protest, which seems to appear at an earlier age in girls than in boys, is the negation, "No!" "Don't!" "I don't want to." "No more." At about the same age the specific protest, "That hurts" and the "All done?" query that perhaps marks the beginning of patient forbearance, begins to be used.

Next in order of development and in degree of adaptiveness is the begging protest, wherein the child makes a specific suggestion that, if carried out, would extricate him from the distasteful situation. He demands to get down from the examining table or to have his clothes put on again, or he asks to go back to the playroom. A little later to develop and somewhat more adaptive in that it invests the protest with a little subtlety and scheming, is the excuse. One of the earliest and commonest excuses is the toilet request "I want to wet again!" Children who have recently undergone the process of training in bladder control have learned that this is an appeal to which the adult usually yields with promptness. At later ages the child sometimes invents a scapegoat in the hope of having his excuse accepted, as in the statement, "My mother said not to take my shoes off."

A note of belligerence is injected into the protests of both boys and girls at 3 years. The child couches his protests in terms indicating that he recognizes the social situation which confronts him, and that he considers the examiner as his adversary. He threatens, "I'm not going to get on the table. If I do, I'll cry." "You are not going to take my shoes off." "Don't talk to me!" Up to this point, however, the child's protests are all aimed at extricating himself from the examination. His protests show no evidence of knowledge or anticipation of specific procedures that are to be carried out; he expresses merely a vague dread. Definite worries begin to appear at 3 years in girls and 3½ years in boys. Both anticipation and specific knowledge of the examinations are evident in such protests as "Don't stick the needle in me!" "I don't want to be measured." The ages at which children begin to show this anticipatory dread of the examinations correspond to that at which they show the greatest amount of emotional upset over the examinations.

Stating the resolve, "I am not going to cry!" makes its first appearance in the protests of both sexes at the age of 3½ years.[4] These brave words, in which the child probably was coached before he left home, are reiterated at later ages with and without variations, such as "The tears keep coming, but I keep wiping them away"; "It didn't hurt, but I don't like it."

Resignation and acceptance of the inevitable is voiced at 5 years in such mild protests as "You are going to make me have my pictures taken"; "Am I going to have my fingers pricked?" "I was choking; is that all?" The final development in the protest is the modified protest; the child says, in effect, "This far you may go with my permission, but pray go no further." He says "Don't hurt me"; "Don't open my mouth too wide"; "Don't put it in too long"; "Don't pull my loose tooth."

At 7 years verbal protests reverted again to the nonspecific outcry, "Oh", "Yow", "Ouch". These are as non-adaptive as the 2-year-old's

[4] For an interesting discussion of the "I don't cry" boast of children and its cultural implications see Lois B. Murphy, Social Behavior and Child Personality. New York, Columbia University Press, 1937, p. 230 ff.

cries for mother, but they represent a further step toward the restrained expletive wherewith our culture permits the adult to give vent to his feeling of pain.

Some of the sex differences seem worthy of comment. In general, girls progressed to the more adaptive level of protests about 6 months in advance of boys; this was particularly true of their remarks indicating acticipatory dread. On the other hand boys gave frequent verbalization of effort at self-control. This was particularly true at the finger-prick situation. At 3½ years one boy said the prick "hurt a little bit; not much, just a little bit." This type of response appeared again at 4½ years in the boast, "It hurts, but I didn't cry," and reached its climax at 6 years in the expression, "Anybody that cries is a sissy!" Girls did not make the "I didn't cry" boast at the blood test until 5 years. In the 4 to 5-year period girls appealed to the nursery assistant for comfort, saying "Stay with me!" "Let me sit on your lap," and "Hold my hand!" This paralleling of expressions of bravery on the part of boys with expressions of appealance on the part of girls suggests the impact of different culture patterns on the two sexes. A little boy must not be a sissy; a little girl may seek help and protection. To be sure, girls do develop the "I don't cry" response; indeed they show it quite as early as boys at the other examinations. The possibility that boys are schooled earlier in the brave acceptance of minor injuries that involve blood-letting is an enticing problem for minor research on our culturation.[5]

Thus the childish protest runs a developmental course from non-adaptive appeal for parental aid to self-reliant negativism, to specific suggestions for getting out of the situation, to veiled excuses, to threats and demands, to anticipatory and modified protests about specific procedures, to resolute attempts at compliance with the social code of not crying, to resignation and acceptance of the situation, to the non-specific and almost involuntary outcry. Although the evidence does not warrant the statement that this hierarchy of response is culturally imposed upon the child, it is certainly true that the child's progress is toward more socially approved expressions of his protests.

REFERENCE

(1) Shirley, Mary M. Common content in the speech of preschool children. Child Develop., 1938, 9, 333-346.

[5] *Op. cit., footnote 4.*

A RELIABILITY STUDY OF SPEECH SOUNDS OBSERVED IN THE CRYING OF NEWBORN INFANTS[1]

ORVIS C. IRWIN AND HAN PIAO CHEN

INTRODUCTION

In any research program the primary concern is to safeguard the data. The term "safeguarding the data" here is taken to mean that they are: 1) reliable and 2) valid. This study is concerned with the first of these two problems. Reliability usually is treated under two headings: 1) the problem of observer reliability and 2) the reliability of the data. The purpose of this study is to investigate both of these problems in regard to early speech sounds, especially when they occur during the crying of newborn infants.

The general status of the problem of the reliability of infant speech sound data has been reviewed by Irwin (1). A study of Irwin and Curry (2) reported on the reliability of recording vowel sounds of newborn infants, and Irwin and Krehbiel (3) in an unpublished study reported both observer and data reliability of transcriptions of vowel and consonants of four-, five-, and six-month infants. In all the studies now being conducted the International Phonetic Alphabet is used to transcribe sounds.

In the present study two experiments were conducted, one of which involved a small sample of crying sounds from each of 40 infants;[2] the other, a fairly large sample from each of 5 infants. The unit of observation was the respiration (2). In the first experiment the sample consisted of the sounds carried on each of 30 exhilations giving a total of 1200 breaths for the group. The total number of individual vowel elements on these 1200 respirations amounted to 1520. In the second experiment the sample consisted of the sounds carried on 180 breaths for each baby giving a total of 900 breaths for the group. The total number of separate vowel sounds amounted to 1080. In addition a further analysis of the Irwin-Curry data has been made. This sample consists of the sounds carried on 25 respirations of each of 40 infants or a total of 1000 breaths which yielded a total of 1285 vowel sounds. The data under analyses then consist of three samples with a grand total of 3885 vowel sounds.

WHINING AND CRYING

In the study by Irwin and Curry a discrimination was made between crying and non-crying sounds. However since non-crying sounds are of rare occurrence in newborns, no report can be made at present concerning them. In the two experiments of the present study a further discrimination was made between crying and whining sounds. A report on the latter will be deferred until later but it does involve the necessity of

[1]*From the Iowa Child Welfare Research Station, University of Iowa, Iowa City, Iowa.*

[2]*Ten of these subjects were negro infants. A systematic study of racial differences in speech sound equipment of newborns will be reported later.*

establishing observer reliability for discriminating these two categories. Accordingly this problem also was undertaken.

Two observers simultaneously but independently recorded 60 breaths on each of 5 infants. Speech sounds as such were disregarded and the record on each of the 300 breaths consisted either of the term "whining" or "crying." In determining the difference between whining and crying in newborns, the observers used the following set of criteria:

1. Regularity of respiration
2. Shape of mouth opening
3. Tongue position
4. Face and lid contraction
5. Loudness of sound

On the basis of these criteria the condition of crying is defined as regularity of breathing, the mouth is opened wide, usually in a rectangular shape, the tip of the tongue is elevated, the muscles of the face are strongly contracted and the eyelids tightly closed, and the sound is uttered with force or loudness. Whining is defined as occurring during irregular breathing, the mouth is partially opened, the tongue tip is not elevated, the muscles of the face are only slightly contracted, the lids are open or if closed are not tightly contracted, and the sounds are feeble.

The following tabulation presents the data on the 5 cases. That crying and whining may be reliably discriminated is evident. The per cent of agreement between the two observers ranged from 95 to 100, the average being 97 per cent.

Subject	Per Cent Agreement
1	97
2	95
3	100
4	97
5	95
Average	97

RELIABILITY IN TERMS OF FREQUENCY AND OF THE OBSERVED BEHAVIOR UNIT

The problem of observer reliability of speech sound data in newborns may be approached in two ways: 1) agreement in terms of the total number of speech elements considered as separate events, and 2) agreement in terms of the pattern of elements occurring on the respiration or breath. The former will be called frequency agreement, the latter breath agreement. In the case of the former the frequencies of each vowel element are calculated as a per cent of the total number of elements uttered regardless of patterning. This is done for each observer and the corresponding values for both are compared on a per cent basis.

A more rigorous analysis is concerned with the agreement between two observers' transcriptions when the pattern of sounds occurring on the breath is considered. The problem will be clarified by indicating several possible cases. The first case is presented when one observer hears the sound ʌ on a breath whereas the other records the sound ɛ.

This constitutes disagreement. In the second case one observer might hear the sounds ξ, and I on a given breath and the second observer also might hear these sounds. Obviously this constitutes 100 per cent agreement. However there is a third possibility in which there is agreement on some of the sounds occurring on a breath and disagreement on others. One observer may record the sounds ξ, Λ on a given breath, whereas the other hears χ, Λ. In this case, of a total of four sounds, two are disagreements and two are agreements. Accordingly the index is 50 per cent. This has been termed the mixed category.

All sounds, therefore, have been analyzed first into three categories: agreements, disagreements, and mixtures. Then the sounds classified in the mixed group in turn have been calculated in per cents of agreement and disagreement, and their values have been added to the original categories of agreement and disagreement. The three categories thus have been reduced to two and in this manner an index has been derived to indicate breath agreement.

EXPERIMENT 1

Table 1 gives the per cent of agreement for two observers for both breath and frequency in experiment 1 on each of 40 babies. The average agreement in terms of the breath unit is 91 per cent. For the frequencies it is 93 per cent.

The following tabulation indicates how the two sets of agreements are distributed. It will be seen that in the case of the breath unit there was agreement above 97 per cent on 12 infants, above 93 per cent on 18 infants, above 85 per cent on 33 of the 40 subjects. In terms of total frequency of sound elements there was agreement above 97 per cent on 17 infants, above 93 per cent with 27 infants, and above 85 per cent with 36 of the 40 subjects. In the case of breath agreement 6 indices

Agreement, Per Cent	Frequency	Cumulative Frequency
Breath Unit		
97 to 100	12	12
93 to 96	6	18
89 to 92	9	27
85 to 88	6	33
81 to 84	1	34
77 to 80	2	36
73 to 76	4	40
Total	40	
Total Frequency		
97 to 100	17	17
93 to 96	10	27
89 to 92	5	32
85 to 88	4	36
81 to 84	2	38
77 to 80	1	39
73 to 76	1	40
Total	40	

TABLE 1

Observer Reliability of Crying Sounds of Forty Babies

Infant Number	Frequency Agreement, Per Cent	Breath Agreement, Per Cent
1	74	76
2	98	74
3	87	77
4	100	100
5	94	74
6	94	86
7	80	88
8	86	87
9	89	79
10	82	89
11	98	96
12	100	93
13	100	100
14	82	76
15	93	89
16	92	82
17	100	100
18	100	98
19	86	85
20	94	98
21	97	99
22	95	91
23	97	99
24	97	90
25	94	96
26	95	88
27	100	96
28	97	97
29	100	100
30	87	90
31	95	89
32	92	88
33	90	91
34	89	92
35	93	91
36	100	94
37	97	94
38	100	100
39	98	97
40	95	97
Average	93	91

fall between 70 and 80 per cent whereas in the case of frequency agreement 2 indices fall between these limits.

Thus with either method a quite satisfactory observer reliability is established.

A graphic method of presenting the dispersion of speech elements involving the principle of the profile was described in a previous report (4). Figure 1 includes six profiles comparing the dispersion vowel sounds of newborns as recorded by three pairs of observers. The two upper profiles give the percent of frequency of observer I and observer

OBSERVED RELIABILITY ON PER CENT OF FREQUENCY OF VOWEL SOUNDS OF NEWBORNS

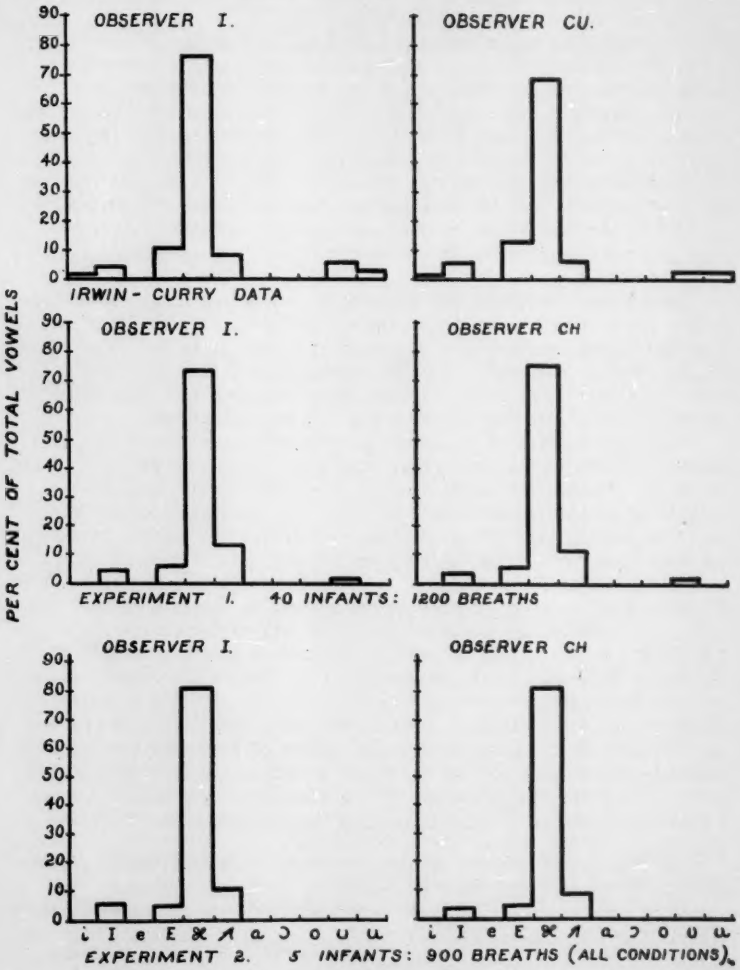

PER CENT OF TOTAL VOWELS
OBSERVER I.
OBSERVER CU.
IRWIN - CURRY DATA
OBSERVER I.
OBSERVER CH
EXPERIMENT 1. 40 INFANTS: 1200 BREATHS
OBSERVER I.
OBSERVER CH
i I e E æ ʌ a ɔ o ʊ u ɯ
EXPERIMENT 2. 5 INFANTS: 900 BREATHS (ALL CONDITIONS)

FIGURE 1

Cu. There is striking similarity between the patterns of these two observers. The two middle profiles were constructed from the data of experiment 1, and likewise show similarity of pattern. The two lower profiles are based on the data in experiment 2. Here again there is great similarity. The comparisons between the records of pairs of profiles quite adequately indicate the degree of observer reliability which may be achieved with this type of data.

It is of interest to note the trends which occur during the course of collecting the data. They are indicated in Figure 2. The curves represent the average agreements of eight successive groups of 5 babies. Thus the observer reliability on the first 5 babies is about 80 per cent. On the last 5 it is about 95 per cent. One curve represents the Irwin-Curry data, the other the Irwin-Chen data. While there are some differences between the two curves, they start together, they end together, and they indicate that an untrained observer will raise his reliability to over 90 per cent after working with about 18 infants. The figure thus is of practical value in conducting research on sounds for it tells approximately how many subjects should be used to train new observers.

There is a special problem involved in the derivation of the reliability index by the two methods. As was noted the per cent of breath agreement is 91, the per cent of frequency agreement is 93. Superficially it would seem that these two values are sufficiently close so that one method could be substituted for the other. That this is not the case will become apparent from the following reasoning.

When the two sets of values are correlated the coefficient of correlation is .75 $\pm$ .03. Ordinarily this might be considered a high correlation. However the coefficient of alienation, which measures the absence of relationship, amounts to .66. This value is too high to permit a substitution of one of these methods for the other. It may be inferred then that the two methods may not rank the infants alike. To test this inference, the index of forecasting efficiency may be applied. In order that one might predict scores derived by one of these methods from scores derived by the other method, a very high correlation is necessary. A correlation of .995 is required to reduce errors in prediction to 90 per cent. A correlation of 75 gives a percentage reduction over chance in errors of prediction of only 32.9. The two methods, therefore, do not rank the infants in the same order. This is another way of saying that the scores from the method of frequency agreement cannot be substituted for those derived by the method of breath agreement. A cautious interpretation of the correlation value of .75 therefore requires that both methods be used in computing observer reliabilities on speech data at the newborn level.

The reason that the two methods do not rank the individuals alike may lie in the fact that they represent two different problems. The method of breath agreement answers the question - what in terms of the observation unit (the single respiration) actually used is the observer agreement. The frequency method answers the question - what agreement may be achieved with a large amount of data collected over an extended period without regard for the observation unit. Quite evidently these are two different problems.

The Irwin-Curry data may be similarly interpreted. The correlation

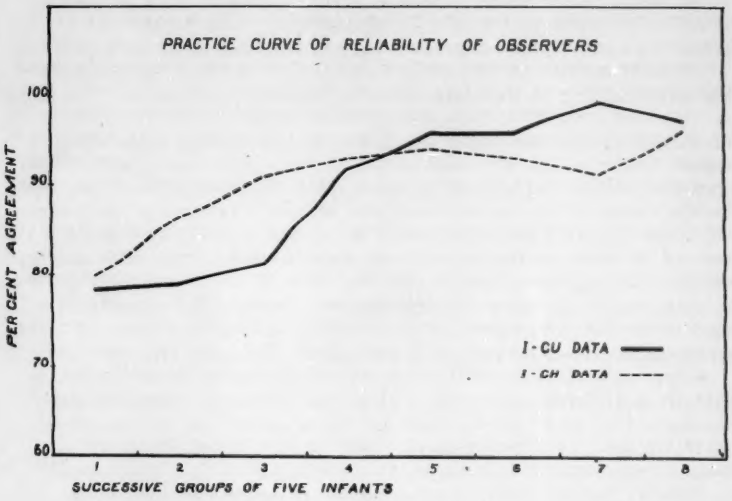


FIGURE 2

value between the breath and frequency ratings of this study amounts to .94. This is much higher than the previous value of .75. Nevertheless the index of forecasting efficiency for this correlation value is 65.8, not a very satisfactory value. The coefficient of alienation is .342 indicating the amount of absence of relationship. Here, too, evidence is forthcoming that the frequency method of establishing observer reliability cannot be substituted for breath reliability. The conclusion that both methods should be employed seems to be re-enforced by this study.

There is a fundamental criticism which may be leveled against this procedure of establishing observer reliability. It is that the untrained observer learns not only the correct habits of transcription from the trained observer but also he learns the errors of the trained observer. This means that the reliability index of 93 per cent is as high as it is because both correct and incorrect habits of transcribing are compared.

The situation has been safeguarded by requiring the trained observer (I) to take all courses in phonetics, both undergraduate and graduate, offered in the university. His original training was conducted by a trained phonetician. In addition he worked with several other experienced phoneticians so that his habits were not determined by a single individual. Moreover, in order to maintain a high degree of accuracy in the use of phonetic symbols, this observer has made it a policy to repeat attendance on these courses and engages in regular practice to maintain skill.

EXPERIMENT 2

In the second experiment 5 babies were used with a total of 180 observations on each. Thus it affords a comparison of a large number of observations on a small sample with the result of the first experiment in which 30 observations were made on each of 40 babies. In addition, this experiment was designed to answer an important question concerning the collection of speech data. In an earlier study (2) it was pointed out that the use of the short time sample failed to yield a satisfactory reliability. Accordingly a short behavior unit, the single breath carrying from one to five sounds, which falls well within the ordinary attention span of the observer, was used instead of a time sample. The individual respirations themselves, however, during violent crying occasionally may occur so rapidly that the observer cannot transcribe the sounds occurring on consecutive breaths. This is a fundamental criticism of the method of the short behavior sample. McCarthy (5) has made this criticism against the general method of phonetic transcription. That it can be circumvented, however, will be apparent from what follows.

A record of 60 breaths was taken simultaneously and independently under three different conditions by the two observers. Transcriptions were made of 60 consecutive breaths, of 60 alternate breaths, and of 60 fourth breaths. If a comparison of these three types of transcriptions should reveal striking differences then the criticism is a valid one. If, however, the per cents of agreement are well above 90 then every second or fourth, etc., breath, instead of consecutive breaths, may be transcribed with satisfactory results. That the latter is the case in terms of breath agreement is evident from the following tabulation which

Infant Number	Condition	Breath Agreement, Per Cent
	Consecutive	85
1	Alternative	84
	Every 4th	84
	Consecutive	97
2	Alternative	98
	Every 4th	97
	Consecutive	96
3	Alternative	96
	Every 4th	97
	Consecutive	97
4	Alternative	98
	Every 4th	98
	Consecutive	97
5	Alternative	98
	Every 4th	96
Average		95
	Consecutive	94
Average	Alternative	95
	Every 4th	94

presents observer reliability of crying sounds of the 5 babies under three conditions. All of the per cent values, it is evident, lie in the nineties except three on the first infant. It will be noted that the average agreements for the three conditions are practically the same.

In terms of frequency agreement the result is the same. This can be seen in the following tabulation.

Infant Number	Condition	Frequency Agreement, Per Cent
1	Consecutive	87
	Alternative	99
	Every 4th	86
2	Consecutive	100
	Alternative	97
	Every 4th	97
3	Consecutive	96
	Alternative	99
	Every 4th	99
4	Consecutive	99
	Alternative	95
	Every 4th	100
5	Consecutive	97
	Alternative	97
	Every 4th	96
Average		96

The average observer agreements under different conditions are:

Consecutive	96
Alternative	97
Every 4th	96

Moreover, so far as observer reliability is concerned, these tabulations reveal that a large sample of sounds from a small sample of subjects gives as satisfactory results as a small sample of sounds from a large sample of infants.

OBSERVER RELIABILITY OF THE FREQUENCY OF SEPARATE VOWELS

Another problem remains to be considered. What is the reliability of the observers on each of the vowel elements? The next tabulation considers this problem for the experiment which involved forty subjects. The second and third columns give the frequency of occurrence of each of five vowel sounds as recorded by the observers. The fourth column shows the per cent of agreement. It will be seen that there is very high agreement, 99 per cent, on the sound æ, which is the most frequent sound. The lowest per cent of agreement occurs in the sound I. Thus

the tabulation shows that for most of the sounds uttered by newborn infants there is good agreement. The average per cent of agreement is 85. When the frequencies themselves are transformed into per cents of the total number of sounds heard, the result is found in the last three columns.

Vowel	Frequency			Per Cent		
	Observer		Agreement, Per Cent	Observer		Agreement, Per Cent
	I	Ch		I	Ch	
I	79	47	60	5.1	3.2	63
ɛ	85	82	96	5.4	5.5	98
æ	1,165	1,153	99	74.7	78.0	98
ʌ	219	187	85	14.0	12.6	90
ʊ	12	10	83	.8	.7	88
Total	1,560	1,479		100.0	100.0	

The following tabulation presents similar analyses of reliability on separate vowels in the second experiment which involved five subjects. Only four sounds were heard during this experiment. The per cent of agreement varies from 82 on the vowel I to 99 on the sound æ. The average per cent agreement here is 89.

Vowel	Frequency			Per Cent		
	Observer		Agreement, Per Cent	Observer		Agreement, Per Cent
	I	Ch		I	Ch	
I	50	41	82	4.6	3.9	85
ɛ	46	52	88	4.2	4.9	86
æ	873	865	99	79.9	81.2	98
ʌ	124	107	86	11.3	10.0	89
Total	1,093	1,065		100.0	100.0	

DATA RELIABILITY

The problem of the reliability of the data will be handled in three ways: 1) by comparing sounds occurring on the consecutive, alternative and fourth breaths; 2) by comparing odd-even items; and 3) by comparing the results of three experiments.

Comparison of Consecutive, Alternative, and Fourth Respirations

In Figure 1 the lower pair of profiles was based upon data collected in experiment 2 under three conditions. These two profiles indicate the results of the three conditions when they are totalled. In Figure 3, however, a further analysis of these data is presented in profiles representing the dispersion of vowel sounds under each of the three conditions. The three profiles arranged vertically at the left show the patterns of the three samples taken consecutively, alternately and on every fourth breath by one of the observers. It is clear that with quite minor variations these three patterns are alike. Likewise the

SAMPLE RELIABILITY OF PER CENT OF FREQUENCY OF VOWEL SOUNDS OF NEWBORNS UNDER THREE CONDITIONS

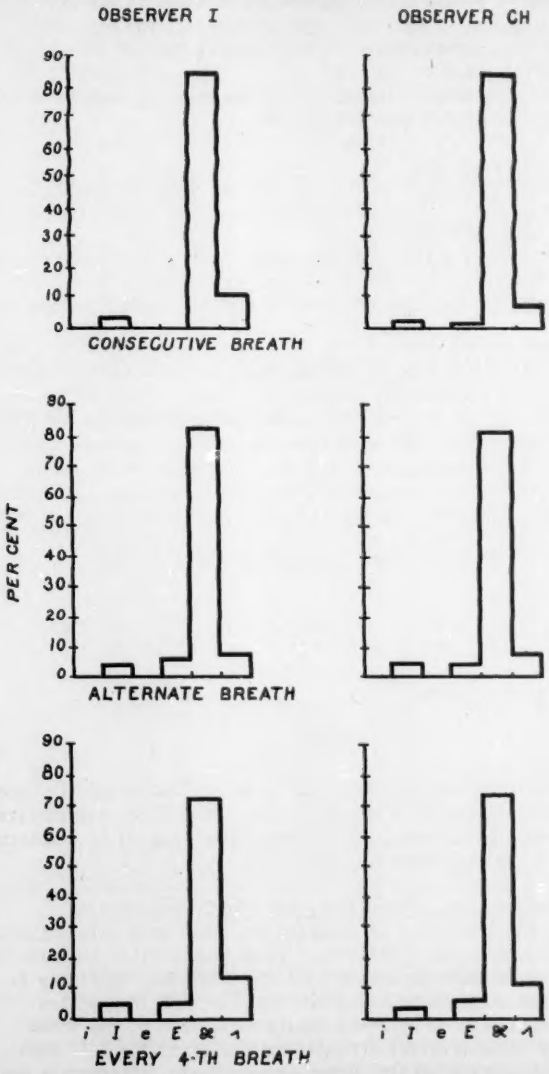


FIGURE 3

three profiles at the right of the figure exhibit a high degree of similarity.

It is readily noted that of the four sounds transcribed in this experiment the vowel æ clearly dominates. Consistently the vowel ʌ maintains second place, while ɪ and ɛ show variation. These vertical comparisons thus demonstrate in a graphical manner the reliability of the data within this sample.

Incidentally a horizontal comparison of the pairs of profiles reveals a high observer reliability within experiment 2.

Comparison of Odd-Even Items

A second demonstration of the reliability of the data is afforded by a comparison of odd-even items. An analysis of these items will be made of the data of both experiments 1 and 2.

The next tabulation presents the odd-even items of different vowels made by the 40 infants in experiment 1. The tabulation is read: for observer I the sound æ occurred 580 times on the odd transcriptions and 585 times on the even, the per cent of agreement being 99. Other sounds are read accordingly.

When the corresponding values for the two observers are averaged and then turned into percentages the results may be seen in the profiles of Figure 4. It is evident from the similarity of these profiles that the reliability of the data in this experiment in terms of odd-even items is quite adequate. In both profiles the æ sound not only dominates the dispersions but their values differ by less than 1 per cent. The order

SAMPLE RELIABILITY OF PER CENT OF FREQUENCY OF VOWELS (40 INFANTS)
EXPERIMENT 1.

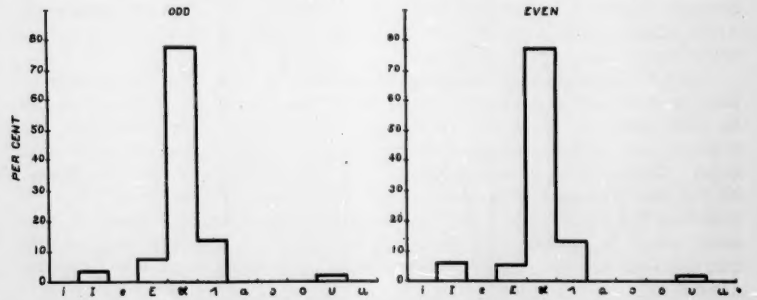


FIGURE 4

of dispersion for the odd and the even values is the same: æ, ʌ, ɛ, ɪ, and ʊ. The corresponding values in each case differ by less than 1 per cent. Separate profiles of the odd-even items of each of the observers have not been included since those of Figure 3, with very slight modifications, could be substituted for them.

In experiment 2, in which five infants were used, the problem of

Vowel	Frequency Odd	Even	Agreement, Per Cent
	Observer I		
I	32	47	67
ɛ	44	41	93
æ	580	585	99
ʌ	109	110	99
ʊ	6	6	100
Total	771	789	
	Observer Ch		
I	19	28	68
ɛ	45	37	82
æ	576	577	100
ʌ	92	95	97
ʊ	6	4	67
Total	738	741	
	Average		
I	25.5	37.5	68
ɛ	44.5	39.0	88
æ	587.0	581.0	99
ʌ	100.5	102.5	98
ʊ	6.0	5.0	83
Total	754.5	765.0	
Average			87

data reliability will be dealt with in two ways: 1) by comparing the odd-even agreement of speech items in each of the categories of consecutive, alternative, and fourth transcriptions; and 2) by comparing these three categories with each other.

Table 2 presents the odd-even frequencies of the different vowels made by five babies under the three different conditions. The table may be read thus: for observer I's data, the vowel sound æ occurred 150 times in his odd consecutive transcriptions, and 151 times in the even ones. The per cent of agreement is 100. This vowel occurred 144 times on his odd alternative transcriptions and 144 on his even ones. It occurred 139 times on his odd fourth transcription and 145 times on his even, etc. The second part of the table gives the corresponding results for observer Ch, and the third part gives the average values for the two observers.

When the averaged results are transformed into per cent values and plotted as profiles the result is seen in Figure 5. The profiles reveal that under the condition of consecutive transcription the order for both the odd and the even values is: æ, ʌ, I, ɛ. For both the alternative and fourth transcriptions the odd and even order is first æ and then ʌ, but there is some variation in regard to the rank of I and ɛ. However the differences between the odd and even values are negligible. The largest difference amounts to 5 per cent between æ recorded on the even transcriptions and æ on the odd transcriptions. This difference is

TABLE 2

FREQUENCIES OF THE DIFFERENT VOWELS MADE BY FIVE BABIES UNDER THE THREE DIFFERENT CONDITIONS

Vowel	Consecutive			Alternative			Fourth		
	Frequency		Agreement, Per Cent	Frequency		Agreement, Per Cent	Frequency		Agreement, Per.cent
	Odd	Even		Odd	Even		Odd	Even	
Observer I									
I	7	9	78	7	5	71	11	11	100
ɛ	0	0		10	10	100	18	8	44
æ	150	151	100	144	144	100	139	145	96
ʌ	24	17	71	16	12	75	29	26	90
Total	181	177		177	171		197	190	
Observer Ch									
I	1	9	11	5	10	50	8	8	100
ɛ	3	5	60	10	9	90	16	9	56
æ	147	147	100	142	145	98	141	143	99
ʌ	12	19	63	20	11	55	24	21	88
Total	163	180		177	175		189	181	
Average									
I	4	9	44	6	7.5	80	9.5	9.5	100
ɛ	1.5	2.5	60	10	9.5	95	17.0	8.5	50
æ	148.5	149	100	143	144.5	99	140.0	144.0	97
ʌ	18	18	100	18	11.5	64	26.5	23.5	89
Total	172	178.5		177	173.0		193.0	185.5	

not significant. Thus here, too, the profiles are similar, indicating that the reliability in terms of the odd-even frequencies within each of the three categories is quite adequate.

Another way of indicating reliability of the data collected on the 5 infants is by a direct comparison of the results of the three conditions themselves. These results are found in the two tabulations and in the profiles.

The first of the two tabulations presents the frequencies of each of the vowels under the conditions of consecutive, alternative and fourth

Vowel	Frequency		
	Consecutive	Alternative	Fourth
Observer I			
I	16	12	22
ɛ	0	20	26
æ	301	288	284
ʌ	41	28	55
Total	358	348	387
Observer Ch			
I	10	15	16
ɛ	8	19	25
æ	294	287	284
ʌ	31	31	45
Total	343	352	370

SAMPLE RELIABILITY OF PER CENT OF FREQUENCY OF SOUNDS OF FIVE NEWBORNS FOR ODD AND EVEN BREATHS UNDER THREE CONDITIONS (BASED ON THE AVERAGED RESULTS OF 1 4 CH)

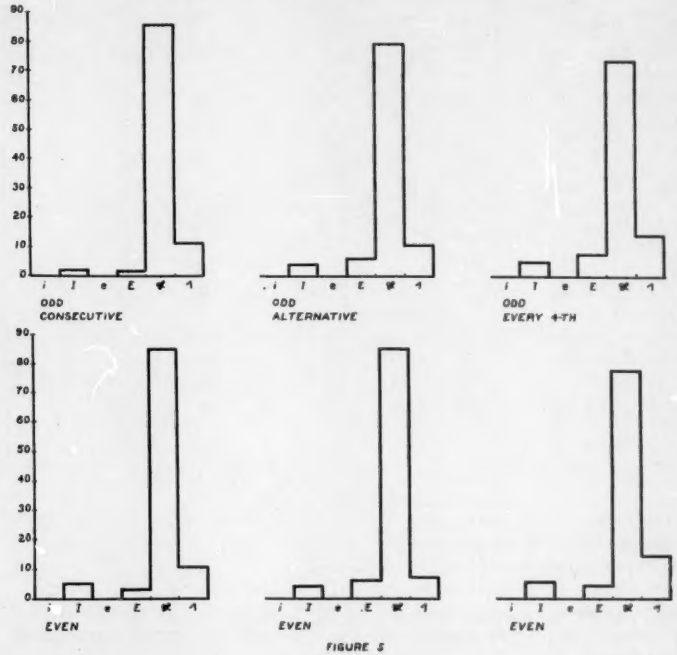


FIGURE 5

transcriptions for the two observers and also for the average of their results. When these values are turned into percentages and these percentages are compared in terms of consecutive vs. alternate, consecutive vs. fourth, and alternative vs. fourth, the results may be seen in

Vowel	Frequency		Fourth
	Consecutive	Alternative Average	
I	13	13.5	19
ɛ	4	19.5	25.5
ɒ	297.5	287.5	284
ʌ	36	29.5	50
Tptal	350.5	350	378.5

the next tabulation. The indices are quite satisfactory on the sound æ, fairly satisfactory for the sounds ʌ, and I, and in two comparisons not satisfactory for the sound ɛ.

Vowel	Agreement, Per Cent: Consecutive vs. Alternative	Consecutive vs. Every 4th	Alternative vs. Every 4th
	Observer I		
I	77	77	60
ɛ			85
æ	98	87	89
ʌ	70	81	57
	Observer Ch		
I	67	67	100
ɛ	43	34	79
æ	95	89	94
ʌ	97	75	72
	Average		
I	95	74	78
ɛ	20	16	84
æ	97	88	91
ʌ	82	78	63

DISTRIBUTION OF VOWEL SOUNDS OF FIVE INFANTS UNDER THREE CONDITIONS

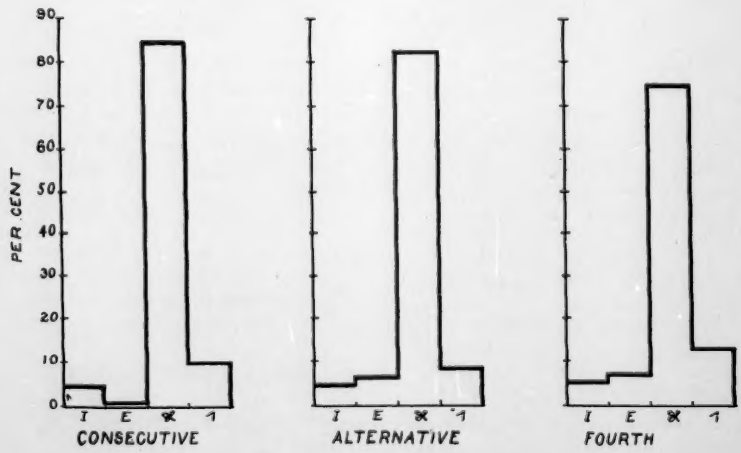


FIGURE 6

When the values in the third part of the frequency tabulation above are turned into per cents and graphed as profiles, the similarity of the patterns becomes evident. The rank order of the sound frequency for ɑ and for ʌ is first and second under all three conditions. For the alternative and the fourth transcriptions ɛ ranks third, and ɪ fourth. However, in the case of the consecutive transcription ɛ and ɪ are reversed.

Comparison of the Results of Three Experiments

In the discussion of observer reliability the similarity between the profile patterns of pairs of observers in Figure 1 was pointed out. Figure 1 also may be used to graphically illustrate data reliability. By reading the figure vertically it will be seen that the successive profiles of the three samples transcribed by observer I are very similar. The same is true of the two samples collected by observer Ch.

In the distribution of the use of the seven vowels shown in the Irwin-Curry data, the sound ɑ is strikingly dominant. The ɛ sound ranks second and a few back vowels are present. Likewise in the dispersion of sounds illustrated in the profiles of experiment 1, the element ɑ is clearly dominant among the five sounds transcribed, and ʌ ranks second. One back vowel is rarely present. Forty infants were used in both of these samples. In the data of experiment 2 on 5 infants, the sound ɑ likewise stands out prominently, the ʌ sound ranks second, but no back vowels were heard.

Thus Figure 1 may be taken as a demonstration not only of observer reliability, but also of data reliability.

SUMMARY

In order to determine the degree of observer and data reliability which can be achieved on speech sound data, two experiments were conducted on the speech sounds of infants under 10 days of age. The first experiment included a sample of vowels occurring on 30 respirations of each of 40 infants; the second, a sample of 180 respirations on each of 5 infants. Transcriptions were made in the International Phonetic Alphabet by two observers working simultaneously and independently. These data together with those of a previous study constitute the materials which have been analyzed. Methods of calculating the reliability of this type of data are described. The following results were obtained:

1. Whining and crying may be reliably discriminated. The agreement index for the two observers averages 97 per cent.
2. Agreement of observers on crying sounds of the group of 40 infants is 93 per cent on a frequency criterion; for a breath criterion, it is 91 per cent. The range in both cases is from 74 to 100 per cent.
3. On the frequency criterion there was observer agreement above 85 per cent in 36 of the 40 infants.
4. On the breath criterion there was agreement above 85 per cent on 33 of the 40 cases.
5. Profiles of vowel dispersion as transcribed by three pairs of observers are strikingly similar.
6. It is evident that training in transcribing infant sounds on about 18 subjects will yield observer reliability above 90 per cent.
7. There are reasons for believing that when the methods of fre-

quency and breath reliability are both used, a more complete result is obtained than when either is used singly.

8. When transcriptions of samples on consecutive, alternative, and fourth sounds are compared as between observers, the frequency reliability is 96 per cent and breath reliability is 95 per cent.

9. For the separate sounds vocalized by the group of 40 babies, the per cent of observer agreement varied from 60 on the sound I to 99 on the sound ɱ. The average for the five vowels uttered by this group is 85 per cent.

10. For the separate vowel sounds of the group of 5 subjects, the per cent of observer agreement varied from 82 on the sound I to 99 on the sound ɱ. The average for four sounds uttered by this group is 89 per cent.

11. When the reliability of the data is analyzed in terms of consecutive, alternative and fourth samplings and presented graphically in profiles of vowel dispersion, a high degree of similarity is evident.

12. A comparison of odd-even items of the vowel sounds of the group of 40 cases shows agreement ranging from 68 to 99 per cent with an average of 87 per cent.

13. A profile comparison of the odd-even items within each of the conditions of consecutive, alternative, and fourth transcriptions shows close similarity.

14. A comparison of the distributions of sounds between consecutive and alternative transcriptions, between consecutive and fourth, and between alternative and fourth transcriptions yields satisfactory reliability on 9 of 12 indices.

15. When the findings of the two experiments of the present study are compared by means of profiles with the results of a previous experiment it is seen that the data show a high degree of similarity.

16. A sample of the sounds of 30 breaths of 40 infants seems to yield more individual sounds than the sample of 180 breaths of 5 subjects.

It is concluded from these findings that quite adequate observer and data reliability can be attained on the speech sounds of newborn infants.

REFERENCES

(1) Irwin, Orvis C.: Research on speech sounds for the first six months of life. Psychol. Bull., 1941, 38, 277-285.

(2) Irwin, Orvis C., and Curry, Thayer: Vowel elements in the crying vocalization of infants under ten days of age. Child Develop., 1941, 12, 99-109.

(3) Irwin, Orvis C., and Krehbiel, Thelma E.: Speech sounds of infants: The fourth, fifth, and sixth months. Unpublished Study.

(4) Irwin, Orvis C.: The profile as a visual device for indicating central tendencies in speech data. Child Develop., 1941, 12, 111-120.

(5) McCarthy, Dorothea: The vocalization of infants. Psychol. Bull., 1929, 26, 625-651.

THE DEGREE OF UNDERPRIVILEGE IN DEPENDENT CHILDREN[1]

RAYMOND R. WILLOUGHBY[2]

The disadvantages suffered by relief children in general comprise an important matter, but one concerning which it has been relatively impracticable to gather any considerable body of information. The existence, however, of a special Federally assisted relief category for children dependent on public funds by reason of the death, incapacity, or absence of parents has made possible a first broad assessment of some of the facts regarding this particular group (for which the usual designation is ADC, for Aid to Dependent Children).

The basic population investigated was the ADC caseload for the city of Providence as of February 15, 1941. These (including some older siblings kept on the records though not actually in receipt of relief) number 615 boys and 612 girls in 341 families, attending 72 public and 18 parochial schools distributed throughout the city. Field work soon showed that certain reductions in this group were necessary; some children had left school, some had not yet gone to school, some were enrolled in special schools, and some were so scattered that the added representativeness would not have justified the time taken to procure their records. No factors are evident which would produce any bias due to these omissions. The final group consisted of 324 boys and 314 girls in 263 families, attending 52 public and 10 parochial schools.

After the project had been cleared with central authorities, assistants were sent into the schools with letters of introduction explaining the purpose and method of making the study; they were instructed to transcribe from the school records (and if these were inadequate from conversations with the teachers or counsellors) five variables on each ADC child and the same five on the next child, in the alphabetical schoolroom list, who was of the same sex and not an ADC child. The assistants reported to the writer after each contact, and the writer was always available to answer questions and make explanations by telephone. The only difficulty of importance encountered was a confusion by some teachers about controls, with consequent disregard of the instructions and assertions that "No other child in the room can be compared with James." This was countered as vigorously as possible, but it may be well to keep in mind, while reading the results, that its occasional occurrence makes the ascertained ADC-control differences minimal.

The variables recorded for each child were: 1) number of siblings; 2) intelligence quotient, from the periodic tests given for record in the Providence school system; 3) marks (so far as they could be readily obtained) for the current and immediately preceding semesters; 4) health notations, usually from the teacher's observation; 5) characteristic disciplinary status, also from the teacher's experience.

Besides the five variables directly under study, three accessory ones are at hand: sex, age, and type of school (public or parochial).

[1]The cordial cooperation of the Providence Schools, public and parochial, is gratefully acknowledged.

[2]From Rhode Island Department of Social Welfare, Providence, Rhode Island.

In the last, the parochial-school children are significantly higher than the public-school children in marks and in percentage noted as being in good health; although there are intelligence records on only a few parochial-school children, careful matching of these with public-school children in the variables at hand does not abolish the difference in marks; i.e. for reasons not connected with intelligence, sex, age, family size, health, or discipline, the parochial children are given higher school marks.

Sex was found to have its usual effects, plus some others. The girls obtain higher marks and have fewer disciplinary notations; they also come from somewhat larger families. The boys have fewer notations of poor health.

In age, the younger children have the advantage in intelligence, marks, and health; there is no relationship between age and discipline, and probably none between age and family size.

Coming to the comparison of ADC and control children in the five principal variables, we find first that the ADC child comes from a larger family. The constants are as follows:

Siblings of	Mean	SD	N
ADC girls	4.15	2.37	301
boys	3.56	2.06	285
Control girls	3.27	2.42	284
boys	3.34	2.33	272

This is not particularly surprising, as with given resources large handicapped families will be forced oftener than small ones to seek public aid.

In intelligence the ADC children have a fairly consistent average disadvantage of about 2 IQ points. By matching pairs for sex, age, family size, health, and discipline this difference is diminished somewhat, but still remains significant. IQ and family size are correlated -0.10 in the ADC group and -0.14 in the controls; both the coefficients and their difference are significant. The ADC children have lower IQ's than the controls for all family sizes except two (in one of which the difference is insignificant). In the unrestricted groups the absolute means are 97.0 and 99.1 for ADC and controls respectively.

For the unrestricted groups, the mean ADC school mark is 75.5 and the mean for the controls 77.8; this is a significant difference. Marks and intelligence correlate 0.34 in the ADC group and 0.43 in the controls, also a significant difference. The ADC children get slightly higher marks then the controls at the lower IQ levels and slightly lower ones at the higher levels, but the stability of this result is somewhat doubtful. The more satisfactory matching procedure (for sex, age, intelligence, family size, health, discipline) gives a mean mark for the ADC's of 75.2 and for the controls of 76.2, which in view of the correlation between the pairs is significant.

The health and discipline notations were transformed into ratings of A, B, and C, the first signifying the most desirable condition; evaluations are made in terms of percentages of A's. These percentages for the health variable are as follows:

Group	Percentage A's
Public ADC's	79
controls	86
Parochial ADC's	85
controls	96
All ADC's	80
controls	87

For the combined schools, the standard error of this difference of 7. per cent between ADC and controls is about 2.2 per cent, and the difference is therefore significant. Triple distributions for marks, health and discipline show clear effects of each of the three variables independent of the others, and the matching technique (for sex, age, family size, intelligence, and discipline) gives 91 of 114 ADC children in A health as against 102 of the corresponding 114 controls -- nearly a 10 per cent difference.

Disciplinary notations show the following differences:

	Percentage A's
ADC	79
Control	85

The relationship to intelligence shows only one reversal:

Discipline, percentage A	IQ	A	B	C+	C	C-	D	E
ADC		86	88	86	81	81	75	59
Controls		98	85	92	85	85	84	73

The matching technique (for sex, age, family size, intelligence, health) still gives a small difference (81 per cent as against 85 per cent) in the same direction.

This investigation fairly definitely indicates, then, that the relief child is handicapped in test intelligence, school marks, health as seen by the teacher, and whatever emotional factors contribute to cooperativeness in school, and that each of these factors has a definite contribution to the total disadvantage. The degree of disadvantage is of the order of about one-fifth that usually found for children of the underprivileged ethnic groups (such as negroes and Indians) living in or on the borders of our culture. The study does not, of course, indicate the cause of the disadvantage; the argument is tenable that there is some biological difference, but it is perhaps more reasonable to suppose that it may be more closely related to the emotional factors commonly found in broken homes.

Contents for December

Popularity among preschool children. *Rosemary Lippitt* 305

The location of the subordinate clause in oral and written language. *Edith A. Davis* 333

Anthropometric and orthopedic examinations—A technique for use with children. *Mabel Pottinger Baum and Vernette S. Vickers* 339

Development and cultural patterning in children's protests. *Mary Shirley and Lillian Poyntz* 347

A reliability study of speech sounds observed in the crying of newborn infants. *Orvis C. Irwin* 351

The degree of underprivilege in dependent children. *Raymond R. Willoughby* 369

Information for Contributors and Subscribers

Child Development, one of three publications issued by the Society for Research in Child Development, is devoted to brief papers of current research in the child field. The journal is issued quarterly in March, June, September and December. The subscription price is $4.00 per year, postpaid. A limited supply of back issues is available at $1.25 per number. Until the supply is exhausted, complete sets, Volumes 1 to 11 inclusive, may be obtained for $22.00.

Other publications of the Society include *Child Development Abstracts and Bibliography* and *Monographs of the Society for Research in Child Development.*

Child Development Abstracts and Bibliography, now in its fifteenth yearly volume, is issued in February, April, June and August, October, December. The subscription price is $5.00 per year, postpaid.

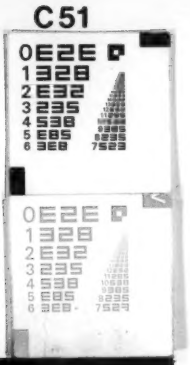
Monographs of the Society for Research in Child Development, issued irregularly during the year, consist primarily of detailed critical reviews and seriatim studies. The subscription price is $4.00 per yearly volume.

The three publications of the Society are available at the special price of $10.00 per year.

Correspondence relating to articles and papers for publication in *Child Development* should be sent to Dr. Carroll E. Palmer, Society for Research in Child Development, National Research Council, 2101 Constitution Avenue, Washington, D. C.

Correspondence relating to studies for publication in the Monograph Series should be sent to Dr. L. W. Sontag, The Samuel S. Fels Research Institute, Antioch College, Yellow Springs, Ohio.

Correspondence relating to subscriptions and business matters for all publications should be sent to the Editorial Office, Society for Research in Child Development, National Research Council, 2101 Constitution Avenue, Washington, D. C.


C51